

# Системы Services Engine на платформах ENVI и IDL: достижение максимальной эффективности

## ВВЕДЕНИЕ

Вот уже 30 лет компания Exelis Visual Information Solutions (VIS) предлагает первоклассные средства анализа научных данных и решений для исследователей, аналитиков и предпринимателей на платформах IDL и ENVI. Поддержка и разработка платформ ENVI и IDL осуществляется с целью удовлетворения потребностей широкого круга пользователей с точки зрения аналитических и системных возможностей. По мере того как пользователи разворачивают деятельность в корпоративных сетях и в облачных средах, система Services Engine обрабатывает и анализирует научные данные по запросу и в режиме онлайн на платформах ENVI и IDL в режиме RESTful-сетевых сервисов. Services Engine представляет собой агностическую систему среднего слоя. Система соответствует стандартам открытости, поддерживает многопользовательский режим (в сети, вход с мобильных устройств, с настольного компьютера), полностью настраивается. Как и в любых крупных проектах корпоративных сетей, облачных сред или кластеров, приоритеты в части преимуществ и возможностей по производительности задаются каждый раз по-новому, на них могут влиять архитектурные решения. В этой статье описаны показатели производительности, которые достигаются системой Services Engine на платформах ENVI и IDL, пути достижения максимальных показателей, а также факторы, ограничивающие эффективность.

## ПРОБЛЕМА

Объем и состав текущих задач в сфере обработки научных и промышленных данных настолько велик, что на практике трудно обработать все эти данные даже на самых современных настольных компьютерах. Использование же известных суперкомпьютеров часто затруднительно из-за стоимости и особенностей выделения ресурсов. Однако применение в корпоративных сетях и облачных средах позволяет пользователям получить значительные вычислительные ресурсы на существующих группах настольных компьютеров, в объединенных системах либо в облаках PAYGO и кластерах, таких, как Amazon Web Services.

Выполнение крупномасштабных аналитических проектов и решение задач — это более сложное дело, чем запуск обычного программного кода одновременно на нескольких компьютерах. Необходимо учитывать управление разбиением и рекомбинацией параллельных задач, управление ресурсами, порядок хранения и доступа к данным в сети без необязательного копирования/перемещения информации. Максимальная выгода от выполнения аналитических задач в корпоративных или облачных системах достигается за счет структуризации самих данных и аналитических процедур таким образом, чтобы добиться преимущества перед простым разделением задач для многих сотрудников.

*Статья предоставлена компанией Exelis VIS. Перевод с английского языка*

## СТАНДАРТЫ ДЛЯ СИСТЕМЫ SERVICES ENGINE НА ПЛАТФОРМАХ IDL И ENVI

Компания Exelis VIS выполняет три тестовых программы, задающие ориентир по максимизации выгоды за счет применения системы Services Engine. В системе Services Engine используется концепция Главного узла и Рабочего узла. В компьютерах может быть сколько угодно рабочих узлов, что позволяет оптимизировать конфигурацию в зависимости от числа процессоров и ядер, а также типа выполняемых задач. Один Главный узел обеспечивает управление и распределение для рабочих узлов и задач.

### СЦЕНАРИИ ТЕСТИРОВАНИЯ

#### Задачи

В ходе тестирования реализуются три различные задачи, соответствующие трем фактическим случаям:

**1) Без процесса и без интенсивного использования диска:**

- Задача — управление сетью.
- Выбранная задача: сложение двух чисел.
- Входные данные: два числа.

**2) Файл, ввод/вывод, интенсивно:**

- Задача — «узкое место» диска.
- Выбранная задача: умножение/сдвиг.
- Входные данные: пространственная замена ([0,4409,0,3522]) из набора данных WorldView-2.

**3) Интенсивная загрузка ЦПУ:**

- Задача — высокая нагрузка на процессор.
- Выбранная задача: FFT (БПФ).
- Входные данные: пространственная замена ([0,4409,0,3522]) из набора данных WorldView-2.

#### Данные

Входные данные представляют собой хранимый на месте панхроматический набор объектов WorldView-2 в формате TIFF,

052310279180\_01\_P001\_MUL\10JAN19161609-M1BS-052310279180\_01\_P001.TIF

- Размеры: 8820×7046×8 [BSQ];
- Тип данных: байт;
- Размер: 499573534 байт;
- Тип продукта: базовый 1B;
- Проекция: \*RPC\* WorldView-2 географическая широта/долгота:
  - Данные: WGS-84;
  - Пиксели: 0,000023 x 0,000022 град

#### Аппаратные средства

Главный узел ЦПУ:

- Процессор AMD Opteron 6174;
- 12 ядер x 4 разъема = 48 ядер на 2,2 ГГц;
- RAM 128 Гб.

#### Рабочий узел:

- Процессор Intel Xeon E5-2670;
- 8 ядер x 4 = 32 ядра на 2,6 ГГц;
- RAM 256 Гб.

### ТЕСТОВАЯ ЗАДАЧА НОМЕР ОДИН

При любых приложениях главного/рабочего типа важно, чтобы управление (установка очередности, диспетчерский контроль и выполнение различных операций) не приводило к замедлению решения основной задачи. Первый пример показывает, что IDL и ENVI системы Services Engine не увеличивают издержек на управление по мере увеличения числа выполняемых задач.

В первом тесте использовался один Главный узел и один Рабочий узел, настроенный на работу с четырьмя сотрудниками. Загружалось 16, 64 и 192 (задачи количество сотрудников ×4, 16, 48), по мере выполнения задач фиксировалось соответствующее время. Результаты, отображаемые на рис. 1, показывают, что увеличение числа задач не меняет времени, затрачиваемого на отдельные задачи, то есть наклон кривой практически отсутствует.

## ТЕСТОВАЯ ЗАДАЧА НОМЕР ДВА

Для Главного и Рабочего узлов, снабженных многими ЦПУ со многими ядрами, необходимо знать количество сотрудников, выделяемое на один узел. Ядра могут рассматриваться как независимые, однако они делят память, пространство жесткого диска и коммуникационные ресурсы. Хотя каждое ядро, в принципе, может заниматься одним работником, наилучшие показатели достигаются, если для узла назначено максимальное число сотрудников, исключающее конкуренцию за ресурсы.

Использовались те же аппаратные средства, задачи и данные, как и в первом тесте. При числе сотрудников на рабочий узел от 2 до 4–8 число задач изменялось от 8 до 16–32 (число сотрудников×4).

Для задачи сложения чисел среднее время на одну задачу не изменяется при увеличении числа сотрудников. Эффективное время на задачу (полное затраченное время, деленное на число задач) слегка сокращается по мере увеличения числа сотрудников. Процессы, запущенные сотрудниками, не борются за ресурсы.

Однако при выполнении операции умножения/сдвига среднее затрачиваемое время на задачу заметно увеличивается вместе с увели-

чением числа сотрудников (рис. 2). Сотрудники начинают конкурировать в борьбе за сетевые и дисковые ресурсы, которые характеризуются максимальными скоростями передачи, независимо от приоритетов в управлении или установления очередности. Указанные ограничения влияют на затрачиваемое сотрудниками время выполнения задач, поскольку сотрудники ожидают очереди в доступе. Сотрудники конкурируют в борьбе за время, выделенное ЦПУ, в меньшей степени, чем за право доступа к диску. Важно иметь в виду, что эффективное время на задачу (полное затраченное время, поделенное на количество задач) сокращается по мере увеличения числа сотрудников (рис. 3). Борьба за сетевые и дисковые ресурсы является доминирующим фактором, ограничивающим скорость выполнения операций умножения/сдвига.

Для решения задачи FFT (БПФ) среднее затрачиваемое время на задачу, согласно рис. 2, увеличивается вместе с увеличением числа сотрудников, однако не в такой степени, как для задачи умножения/сдвига. Сотрудники борются за доступ к диску, ЦПУ и RAM, однако не до точки перегрузки памяти. Эффективное время, затрачиваемое на задачу (полное затраченное время, поделенное на число задач), существенно снижается по мере увеличения числа сотрудников (рис. 6). В отдельных задачах происходит борьба за доступ к диску, однако отрезки времени, затраченного на операции в ЦП, существенно превышают задержки из-за конкуренции. Снижение требований к диску, вместе с собственным планировщиком OS, позволяет выполнять каждое задание по FFT (БПФ) независимо и поэтапно, что повышает пропускную способность системы.

Наилучшее соотношение числа сотрудников к числу машин или ядер для задач каждого типа зависит, в основном, от пользовательского сценария и представления задания (одиночные задания или группы). Число сотрудников влияет как на среднее время, затрачиваемое на задачу, так и полное затрачиваемое время на серию задач. Каждая задача может выполняться

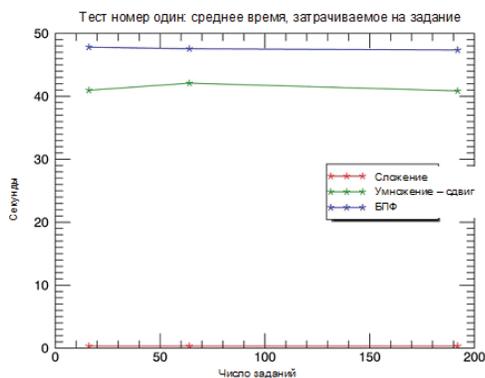


Рис. 1. Среднее затрачиваемое время на задание для теста номер один. Рассчитано как сумма времен, затрачиваемых на отдельные задачи, поделенная на полное число задач

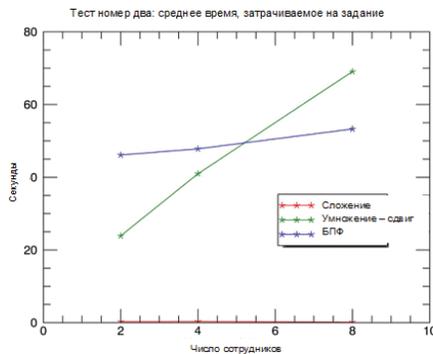


Рис. 2. Среднее затрачиваемое время на задание для теста номер два. Рассчитано как сумма времен, затрачиваемых на отдельные задачи, поделенная на полное число задач

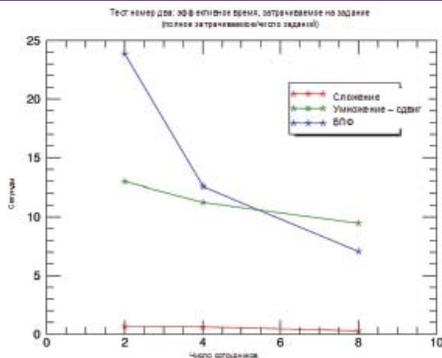


Рис. 3. Среднее затрачиваемое время на задание. Рассчитано как время, затрачиваемое на серию заданий, поделенная на полное число задач

быстрее при меньшем количестве сотрудников. Это особенно верно в том случае, если запущенные задачи борются за ресурсы.

В то же время серии задач (полный поток) выполняются быстрее, если сотрудников больше, вплоть до точки перегрузки. Собственный планировщик операционной системы и изменяющиеся требования к задаче позволяют нескольким задачам выполняться одновременно, без взаимного влияния. В этом случае отдельная задача может выполняться дольше, однако пропускная способность системы фактически увеличивается. Это видно из рис. 2, на котором время, выделенное на отдельные задачи, увеличивается с увеличением числа сотрудников, однако серии задач выполняются быстрее с увеличением числа сотрудников, согласно рис. 3.

Неэффективность наблюдается, если нехватка ресурсов не преодолевается собственным планировщиком ОС и системой распределения. На рис. 4 показан данный сценарий так, как он реализуется на тестовом кластере Exelis VIS. Оптимальная пропускная способность достигается в нижней точке линии, соответствующей четырем сотрудникам на кластере. Если сотрудников не четыре, выполнение задачи FFT (БПФ) существенно замедляется, поскольку задачи, борющиеся за выделение RAM, и операционная система в целом вынуждены сбрасывать данные на диск через систему виртуальной памяти (перегрузка).

### ТЕСТОВАЯ ЗАДАЧА НОМЕР ТРИ

В последнем тесте используются те же Главный узел, задачи и данные, что и в первых двух, но с четырьмя сотрудниками на один рабочий узел и с изменением числа рабочих узлов от 1 до 4. Число задач изменялось от 8 до 16 и 32 (число сотрудников  $\times 4$   $\times$  число узлов). Этот тест показывает, как время выполнения задач системами Services Engine IDL и ENVI изменяется с изменением размера кластера.

Как показано на рис. 4, среднее затрачиваемое время на каждую задачу сохраняется по мере увеличения числа узлов.

На рис. 5 показано полное затрачиваемое на задачу время, приведенное на один узел. Небольшой положительный уклон, особенно по операции сложения, соответствует управлению Главного узла. Этот небольшой уклон показывает, что операции системы Services Engine ENVI и IDL эффективно масштабируются по мере добавления узлов.

Наконец, на рис. 6 показано эффективное время, затрачиваемое на задачи, которое сокращается по мере добавления узлов в кластер, аналогично результатам, приводимым для теста номер два. Это — наихудший сценарий, поскольку все задачи производят запись и чтение в общие зоны/файлы диска, и он в высшей степени нехарактерен для текущей версии приложения.

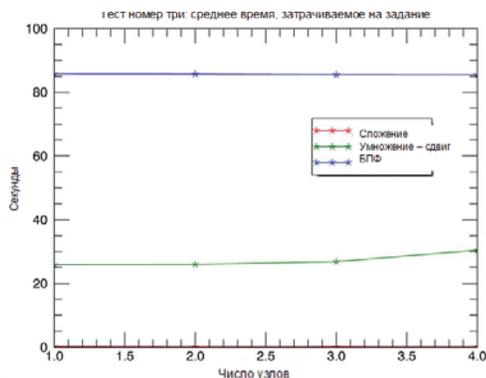


Рис. 4. Среднее затрачиваемое время на задание для теста номер три. Рассчитано как сумма времен, затрачиваемых на отдельные задачи, поделенная на полное число задач

## ВЫВОДЫ: МАКСИМИЗАЦИЯ ПРОИЗВОДИТЕЛЬНОСТИ IDL И ENVI СИСТЕМЫ SERVICES ENGINE

Максимальная производительность системы Service Engine зависит от соответствия конфигурации типам системы обработки. Поскольку эта система как высокопроизводительна, так и масштабируема, Services Engine обеспечивает максимальную загрузку вашей системы. Качественная разработка системы гарантирует, что ни один из компонентов не может нарушить функционирование системы в целом:

- если задача, по сути, вычислительная (например, БПФ по изображению), и оптимизация доступа к данным и ввода/вывода не является существенной проблемой, то не так уж важно, сколько сотрудников подключено к кластеру;
- если система Services Engine используется в стандартной настройке кластера, следует использовать улучшенные режимы доступа при более чем восьми сотрудниками, производящих операции ввода/вывода в интенсивном режиме;
- твердотельные драйверы, системы NAS и оптимизированные массивы RAID помогают ускорить интенсивные операции ввода/вывода настолько, это возможно, если система Services Engines настроена более чем на 8 сотрудников;

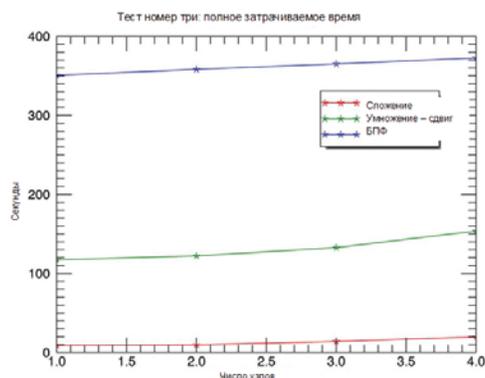


Рис. 5. Полное затрачиваемое время для теста номер три. Рассчитано как время на выполнение серии задач

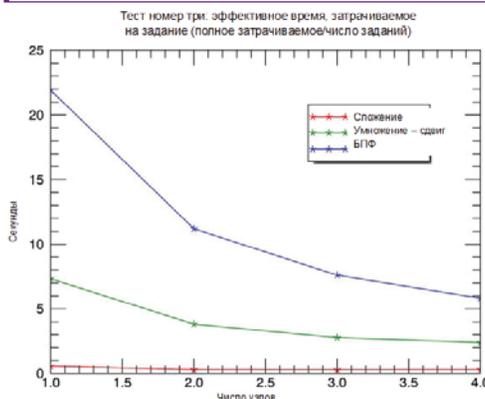


Рис. 6. Эффективное затрачиваемое время на задание для теста номер три. Рассчитано как время на выполнение серии задач, поделенное на полное число задач

- если система настроена на работу с Services Engine, необходимо убедиться, что на каждом узле достаточно памяти для удовлетворения требований по задачам, требующим больших ресурсов памяти, с учетом числа сотрудников на каждом узле.

Система Services Engine на платформах IDL и ENVI представляет собой «мост» между разработкой прототипа для настольного компьютера и доступными высокопроизводительными вычислениями в облачных средах. Ее можно сконфигурировать в соответствии с потребностями пользователя так, чтобы добиться максимальной производительности и эффективности.