



Алексей Жирков, Максим Попов

Исследование средств для работы с Big Data в промышленности

В статье рассказывается о том, как обрабатывать большие данные в АСУ ТП и какие аппаратные ресурсы можно применять для их обработки на промышленном предприятии. Рассматриваются типовые структуры больших данных, и предлагаются решения для работы с такими структурами на базе современных информационных технологий.

Уже сейчас, и тем более в будущем, самым ценным продуктом на рынке будут данные, вернее, сконцентрированная информация, которую можно использовать для оптимизации деятельности и повышения эффективности, но эта информация должна быть определённым образом получена, накоплена, преобразована и представлена в виде, позволяющем формировать стратегию развития компании. Для выполнения этих задач необходимо проводить исследования в области Big Data (рис. 1). Далее мы постараемся ответить на вопрос о том, как и на каких ресурсах можно обрабатывать подобные данные промышленным предприятиям, кото-

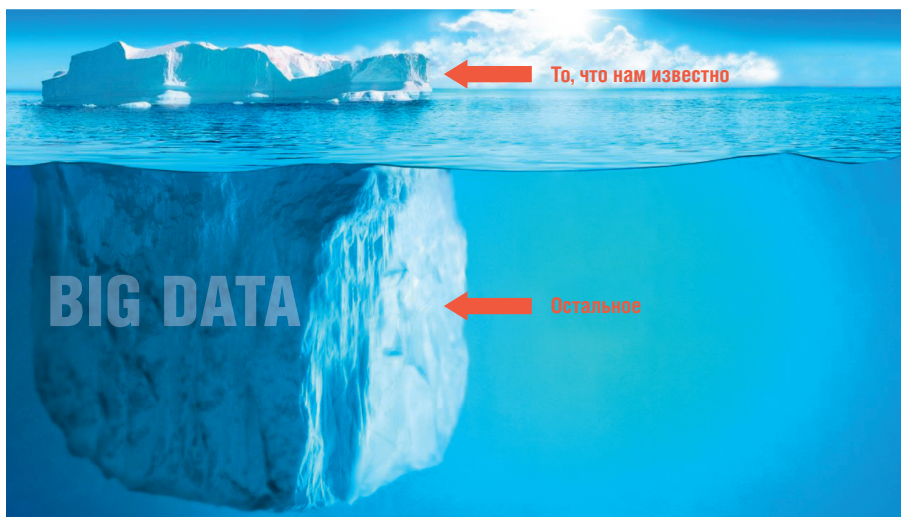
рые сейчас всё активнее начинают использовать методы и средства Big Data в своей деятельности.

КАКИЕ У BIG DATA ПРИЗНАКИ?

Термин Big Data (большие данные) появился не так давно. Бурный рост вычислительных мощностей современной компьютерной техники заставил использовать этот термин, так как возникла необходимость формулировать задачи, в которых речь идёт об обработке структурированных и неструктурированных данных с потоком в сотни тысяч и миллионы параметров в секунду.

Откуда берутся большие данные? Применительно к компаниям, которые занимаются производством и реализацией продуктовых и тем более сервисных решений, сейчас очень важно знать актуальные потребности конечного пользователя. Компании начинают исследовать трафик и запросы в браузерах, в почтовых клиентах, в социальных сетях, анализируют активность использования специализированных приложений для мобильных устройств Интернет-магазинов, контент услуг, различных сервисов, например, заказа такси. Даже возможность изучить такие косвенные показатели, как частота регистрации пользователей в публичной Wi-Fi сети, проанализировать, какие устройства регистрируются в сети, какой объём и класс трафика использует клиент, позволяет учесть все эти данные при формировании будущей стратегии развития информационной сети и средств предоставления контента. И это не говоря о возможностях улучшения общей инфраструктуры, таких как развитие логистики, открытие новых точек продаж или общественного питания, в зависимости от анализа трафика в данной точке.

Все перечисленные источники рожают огромный поток данных, и чем больше секторов рынка и областей интересов пользователей попадает в исследование, тем выше качество прогно-



© Shutterstock / Фотобанк Фотоужения

Рис. 1. Загадка Big Data

зирования и точность аналитических отчётов.

Что касается промышленных предприятий, то основной «виновницей» появления большого количества информации является интеграция. Сейчас на любом промышленном предприятии идут по пути интеграции всех имеющихся систем автоматизации производственных процессов в единую среду. Основное технологическое оборудование, вспомогательные агрегаты, системы собственных нужд и даже административные процессы внутри предприятия становятся источниками потоков данных. В качестве данных здесь фигурируют, конечно же, все технологические параметры, к ним добавляются данные из журналов учёта и эксплуатации оборудования, информация о проведении плановых и внеплановых ремонтных работ, расследовании аварий и происшествий, логистические показатели, сведения о взаимодействии со структурными подразделениями, поставщиками сырья и энергоносителей. Здесь наблюдаем то же положение: чем больше охват и глубина проникновения в функциональную среду предприятия, тем проще будет строить зависимости и находить новые пути оптимизации производства. Но есть на этом пути и нюансы. Например, анализ данных порождает новые данные, и процесс этот может быть почти бесконечным. Также следует принять во внимание, что для качественного анализа данных, кроме самих алгоритмов и программного обеспечения, требуются серьёзные вычислительные ресурсы.

Когда речь идёт о больших данных, то их характеризуют пять основных признаков:

- **Объём данных.** Например, тепловая электростанция с двумя энергоблоками общей мощностью 450 МВт имеет объём 70 000 сигналов ввода-вывода по всем системам автоматизации основного и вспомогательного оборудования, которые, в свою очередь, ежедневно увеличивают объём «сырых» данных на десятки гигабайт. Сигналы от газовых и паровых турбин (рис. 2) формируют относительно небольшой пакет данных, но по причине скорости физических процессов эти данные могут меняться с очень высокой частотой, что также ведёт к очень большому трафику, например, данные виброконтроля оборудования.
- **Скорость передачи данных.** Любое предприятие, производящее техниче-

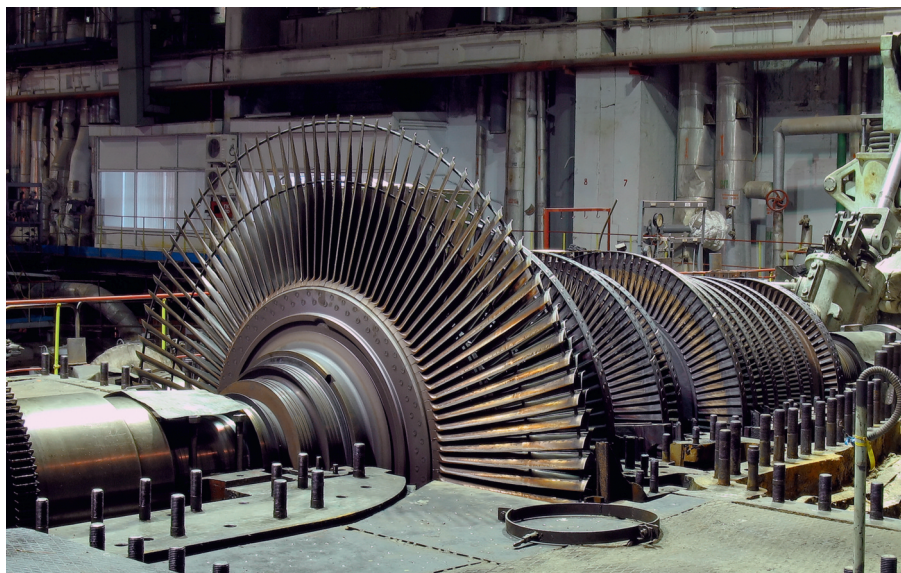


Рис. 2. Паровая турбина

ски сложные изделия, проводит испытания этих изделий, и если взять, например, авиационный двигатель, то при работе на испытательном стенде в течение всего получаса может быть сформирован объём в десятки терабайт информации. Чтобы её передать, требуются надёжные и высокоскоростные каналы передачи данных.

- **Разнообразие данных.** Насколько разнообразен парк оборудования, насколько широко охватывает производство разные технологические отрасли, настолько многообразны параметры, с которыми приходится работать при анализе больших данных. Возвращаясь к тепловым электростанциям, можно отметить необходимость контроля химического состава в процессах водоподготовки, физических свойств теплоносителя и механических свойств оборудования, а также процессов, протекающих в электроустановках преобразования и передачи электрической энергии.
- **Достоверность.** Важно, чтобы получаемые данные имели достаточную точность, от которой зависит правильность принятия решений.
- **Ценность.** За организацией каждого информационного канала стоят определённые затраты, и, конечно, если получаемая информация не имеет никакой пользы, то можно смело считать, что это неэффективное использование ресурсов. Но нужно быть очень осторожным и случайно не избавиться от тех данных, которые косвенно могут принести значительную пользу. И тут на помощь придут алгоритмы, которые могут исследовать скрытые зависимости в уже со-

бранных данных и предоставить оператору возможность определиться с необходимостью дальнейшего использования источников данных.

АРХИТЕКТУРНЫЕ РЕШЕНИЯ ДЛЯ БОЛЬШИХ ДАННЫХ

На первый взгляд, для обработки большого количества данных нужно всего лишь увеличить пропускную способность каналов передачи данных, повысить объём хранилищ и, конечно, нарастить вычислительную мощность процессоров. Но тут возникают множественные ограничения, как технического характера, так и физического уровня. Например, увеличивать частоту работы канала передачи данных выше определённого физического предела не получается, следовательно, разработчики применяют методы распараллеливания потоков данных. То же самое происходит и с делением вычислительной мощности. Но конечный пользователь не должен видеть всей сложности организационной структуры, и на помощь приходит кластерный метод построения систем. Представление вычислительной системы как единого целого значительно упрощает работу пользователя, но повышает сложность инфраструктурных решений. Однако наличие совершенно одинаковых составных элементов в структуре кластера позволяет легко увеличивать количество этих элементов и наращивать тем самым производительность системы в целом.

У пользователя появляется возможность запускать свои задачи в общей среде и динамически выбирать для них выделяемые вычислительные мощности. Например, если существует сервер

обработки данных, который функционирует на двух виртуальных машинах, и возникает необходимость увеличить производительность сервера (возросло число подключаемых пользователей, и, следовательно, повысился объём запросов), на непродолжительное время мы можем легко остановить одну из виртуальных машин и изменить её характеристики (количество ядер или выделяемой памяти), запустить и после успешной синхронизации выполнить такую же процедуру на дублирующем сервере. Причём при использовании кластерной конфигурации этот процесс ещё больше упрощается, ввиду того что добавление ресурсов можно безболезненно производить не только на уровне виртуализации, но и на физическом уровне.

Такой подход наиболее очевиден, но он наглядно показывает, что работа с виртуальной архитектурой значительно сокращает риски и затраты на реализацию многих ответственных и критических процессов.

В виртуальной среде единого пространства могут существовать различные серверы обработки данных и рабочие станции. На практике пользователю также очень важно иметь возможность быстро менять структуру своей системы, перенастраивать её с учётом меняющихся объёмов и типов данных, выделять компоненты, участвующие в ответственных процессах, таких как, например, финансовые операции или операции обеспечения технологического процесса на производстве. Для проведения предварительной разработки интересно наличие тестовых стендов, на которых можно отрабатывать программные модули в безопасном режиме и проводить различные испытания алгоритмов, в том числе нагрузочные испытания вычислительных систем.

Таким образом, для эффективной работы с большими данными нужна очень гибкая платформа, которую можно легко трансформировать, меняя не только конфигурацию, но и производительность узлов.

ПРОЦЕССЫ ПРИ РАБОТЕ С ДАННЫМИ

Для ответа на вопрос: «Как должна быть адаптирована ИТ-архитектура для работы с большими данными?» — прежде всего нужно рассмотреть основные процессы, которые происходят при обращении с массивами информации:

- получение данных;

- первичное преобразование данных, их нормирование, определение целостности и достоверности, присвоение дополнительных маркеров (метка времени, достоверность, актуальность и т.д.);
 - накопление или логирование данных (архивирование);
 - извлечение данных из архива и их математическая и статистическая обработка;
 - предоставление пользователю выходной аналитической информации (графики, отчёты, тренды, таблицы и т.п.);
- Рассмотрим каждый из этих этапов подробнее.

Получение данных. В качестве источников данных могут выступать как полностью автоматизированные устройства сбора данных, так и источники ручного ввода, например, данные исследования лабораторных проб, которые выполняются дежурным персоналом. К любому источнику данных необходимо применять определённые требования:

- данные должны снабжаться меткой времени;
- они должны иметь признак достоверности;
- должна быть обеспечена заданная точность данных;
- информация не должна быть скомпрометирована и должна сохранять целостность;
- помимо основного потока данных (полезная информация), источник должен передавать и служебную информацию, по которой можно оценить и продиагностировать аппаратную и программную часть самого источника данных (датчик, измерительный преобразователь, контроллер, коммуникационные модули).

Первичное преобразование данных, их нормирование — задача, которую лучше всего выполнять как можно ближе к источнику данных и не нагружать этой рутинной центральной систему. Примером такого преобразования может служить пересчёт измеренных значений в другие единицы, например, значение расхода энергоносителя измерительный преобразователь выдаёт в gal/sec (галлон в секунду), и это значение нужно пересчитать в м³/ч и уже в таком виде передавать в центральную систему обработки данных. С подобными задачами легко справляются шлюзы, преобразующие программные и физические протоколы передачи данных, как правило, их производительности хватает на проведение элементарных математических операций.

Отдельная задача — **определение достоверности данных**. Примером является измерительная система, где есть три резервированных датчика, и если два из них выдают примерно одинаковое значение, а третий датчик выдаёт значение, значительно отличающееся по величине или динамике изменения, то можно выявить факт нештатной работы измерительного канала и значения нужно передавать с признаком недостоверности.

Накопление данных — на первый взгляд, очень простая задача. Основными показателями являются скорость записи данных на физический носитель, объём хранилища данных и скорость доступа к информации. Современные системы хранения можно разделить на системы с использованием твердотельных накопителей (solid-state drive, SSD), накопителей на жёстких магнитных дисках (hard magnetic disk drive, HDD), ну, и до сих пор актуальными считаются хранилища с использованием магнитной ленты, когда нужно обеспечивать длительное хранение большого объёма, порядка десятков и сотен петабайт, конечно, с оговоркой о том, что доступ к этим данным не критичен по времени. При таком разнообразии технологий для современной системы хранения данных оптимальным решением будет гибридная архитектура, в которой сочетаются все три технологии, но доля каждой из них выбирается для конкретного случая.

Самая сложная и «процессороемкая» задача — это **обработка данных и получение нового информационного материала** — расчётов, отчётов и аналитических выводов. Для данного уровня важно уметь не только быстро считать, но и очень быстро обмениваться данными с хранилищем. Тут можно предложить использовать технологии высокопроизводительных сетей.

Если ещё нет чёткого понимания, на основе какого программного обеспечения и на каких ресурсах строить собственный центр обработки больших данных, компании чаще всего предварительно проводят исследование и апробацию новых технологий на тестовых площадках или на площадках интеграторов систем, предоставляя им исходное задание. В любом случае, прежде чем останавливать свой выбор на той или иной системе, заказчик должен получить достаточно объективные данные по производительности будущей системы, наглядно убедиться в том, как и каким образом реализованы те или иные функции.

Для решения таких задач оптимально подходит организация испытаний системы на базе компактного испытательного стенда. Естественно, заказчик, проводя подобные испытания, хочет сравнить несколько систем и выбрать лучшую, и для этой цели он ставит общие ограничения всем участникам «соревнований». Как и на олимпийских играх, спортивные снаряды строго стандартизированы, а использование допинга запрещено. Ответственность за соблюдение условий лежит, конечно же, на совести самих организаторов, так как в конечной реализации, если вдруг не получится задействовать «тайную кнопку» ввиду её отсутствия, могут начаться разбирательства, чтобы выяснить, каким образом были получены высокие результаты на испытаниях.

Конечно, может рассматриваться вариант аренды вычислительных мощностей, но если предприятие уже сейчас готово вкладывать средства в реальное оборудование, то лучше сразу обратить внимание на масштабируемые системы, которые могут легко расти параллельно с запросами.

Далее описывается конкретный пример тестового проекта, реализованный компанией ПРОСОФТ.

ПРОЕКТ СИСТЕМЫ ОБРАБОТКИ И ХРАНЕНИЯ ТЕХНОЛОГИЧЕСКИХ ДАННЫХ

На основе ресурсной базы компании ПРОСОФТ требовалось провести моделирование, анализ и испытания системы обработки и хранения технологических данных (СОХТД).

Назначение СОХТД – обеспечение информационных систем и специалистов предприятия оперативными и историческими консолидированными данными по технологическим и производственным процессам. Основными требованиями к функциональным возможностям такой СОХТД являлись:

- исключение дублирования решений по сбору и подъёму данных с уровня предприятия и обеспечение единой точки доступа к оперативным и историческим данным по технологическим и производственным процессам;
- обеспечение реализации эффективных методов обработки и хранения больших объёмов исходных данных;
- обеспечение средств алгоритмической обработки данных и анализ временных рядов с целью информационной поддержки пользовательских аналитических задач.

СОХТД должна обеспечить поддержку данными и аналитическими средствами следующих процессов и задач:

- оперативный учёт продукции;
- планирование и мониторинг грузопотоков продукции;
- планирование ремонта и замены оборудования;
- расчёт наработки оборудования;
- анализ сообщений и формирование событий по инцидентам;
- определение отказов оборудования;
- анализ работы аналоговых датчиков;
- исследование надёжности системы транспортировки продукции;
- анализ КПД агрегатов для собственных нужд;
- формирование диспетчерской отчётности.

Вот что требовалось получить по итогам моделирования:

- выработать приближённую модель данных для дальнейшего исследования;
- определить набор базовых процессов и запросов для оценки производительности системы;
- разработать средство оптимизации алгоритмов обработки данных;



DO-178



EN 50128



Средства разработки программного обеспечения критически важных для безопасности сертифицируемых встраиваемых компьютерных систем

WWW.AVDSYS.RU

ООО «АВД Системы», тел: (916) 194-4271, email: avdsys@aha.ru

IEC 61508



ISO 26262



IEC 62304



Реклама

- получить инструмент для выработки стандартных решений для типовых задач анализа и представления данных;
- собрать данные для оценки оптимального выбора конфигурации вычислительной системы.

Стенд для испытаний

По сути, итогом моделирования становится создание системы автоматизированного управления технологическими процессами в топливно-энергетическом комплексе с функциями учёта и обработки больших данных. Такая система в несколько раз повышает эффективность работы предприятия и надёжность производственного процесса.

Тестирование проводилось на офисном суперкомпьютере с непосредственным водяным охлаждением – Eurotech Aurora G-Station (рис. 3). Задействованная в работе станция была укомплектована четырьмя вычислительными узлами, каждый из которых содержал два процессора Intel Xeon E5-2650 v2, два GPGPU-видеоускорителя NVIDIA K20 (в тестировании не использовались), 64 Гбайт оперативной памяти DDR3 и твердотельный накопитель ёмкостью 240 Гбайт, при этом оперативная па-



Рис. 3. Aurora G-station

мять организована в два NUMA-узла по 32 Гбайт каждый. Все вычислительные узлы объединены высокоскоростной коммутационной сетью Mellanox QDR ConnectX-2 InfiniBand (IB), а также стандартной сетью Ethernet 1 Гбит/с. Пропускная способность сети IB составляет 3,1 Гбайт/с, тогда как соответствующий параметр для сети Ethernet 1 Гбит/с – всего лишь 120 Мбайт/с.

Кэш данных и инструкций в Intel Xeon E5-2650 v2 организован следующим образом: один общий на процессор кэш третьего уровня (L3) объёмом 20 Мбайт и независимые кэш первого и второго (L1/L2) уровней на каждом из ядер процессора. Суммарное вычислительное поле было представлено 8 процессорами с 64 ядрами архитектуры Intel Ivy Bridge с тактовой частотой 2,6 ГГц каждое. Согласно результатам теста производительности High-Performance Linpack (HPL), пиковая (теоретическая) производительность данной конфигурации составила 1,2 (1,3) Тфлопс. В процессе тестирования G-Station работала под управлением операционной системы Microsoft Windows Server 2012 R2.

Расчёты проводились с системными настройками по умолчанию, а именно:

- отключение Hyper-Threading (HT), то есть число логических процессоров было равно числу физических ядер;
- технология динамического изменения тактовой частоты ядер процессора Intel® Turbo Boost была активирована; с активированным Intel Turbo Boost в состоянии простоя тактовая частота всех ядер Intel Xeon E5-2650



ПРОИЗВОДСТВО ЭЛЕКТРОНИКИ ОТВЕТСТВЕННОГО ПРИМЕНЕНИЯ



КОНТРАКТНОЕ ПРОИЗВОДСТВО

Контрактная сборка электронного оборудования

- ОКР, технологические консультации
- Макеты, установочные партии
- Полное комплектование производства, поддержание складов
- Серийное плановое производство
- Гарантийный и постгарантийный сервис

ЗАКАЗНЫЕ РАЗРАБОТКИ

Разработка электронного оборудования по ТЗ заказчика в кратчайшие сроки

- Модификация КД существующего изделия
- Разработка спецвычислителя на базе COM-модуля
- Конфигурирование модульного корпусированного изделия
- Сборка магистрально-модульной системы по спецификации заказчика
- Разработка изделия с нуля

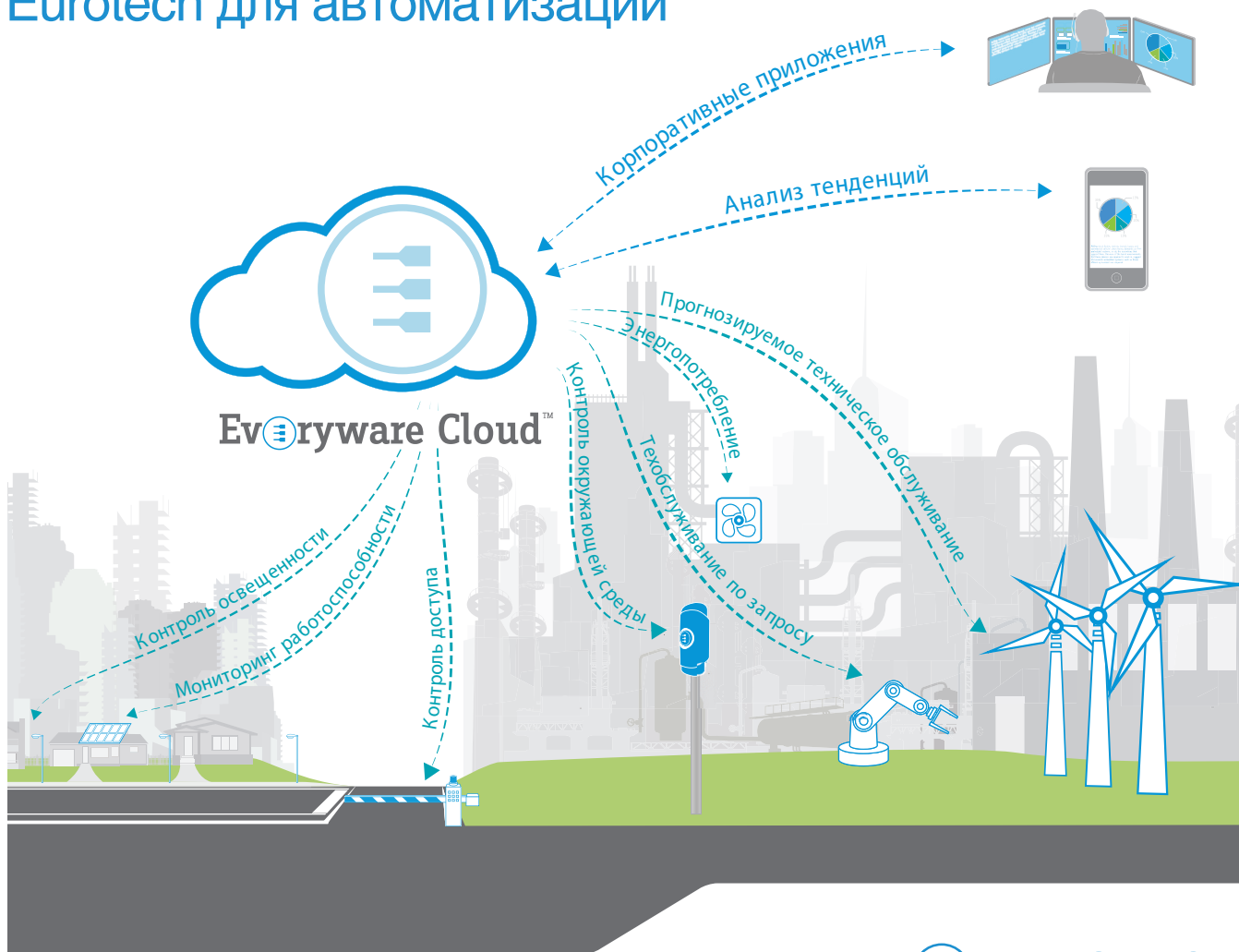
• Опыт работы со спецсчетами и в рамках конкурсных процедур •

ТЕЛ.: (495) 739-0775

WWW.DOLOMANT.RU

Реклама

Облачные технологии Eurotech для автоматизации



Решения Eurotech позволяют заказчикам удобно и безопасно подключать оборудование и датчики к корпоративным программным приложениям с помощью **Everyware Cloud™** — M2M-платформы.

Выполняемые функции

- Управление устройством
- Приложение для устройства и управления жизненным циклом
- Контроль состояния устройства/связи в режиме реального времени
- Поддержка промышленных протоколов
- Простая интеграция с корпоративными приложениями
- Сбор потоков данных с различных устройств в реальном времени
- Анализ данных в реальном времени, их хранение и предоставление исторических данных



ОФИЦИАЛЬНЫЙ ДИСТРИБЬЮТОР ПРОДУКЦИИ EUROTECH

| | |
|---------------------|--|
| МОСКВА | Тел.: (495) 234-0636 • Факс: (495) 234-0640 • info@prosoft.ru • www.prosoft.ru |
| С.-ПЕТЕРБУРГ | Тел.: (812) 448-0444 • Факс: (812) 448-0339 • info@spb.prosoft.ru • www.prosoft.ru |
| АЛМА-АТА | Тел.: (727) 220-7140/7141 • sales@kz.prosoft.ru • www.prosoft-kz.com |
| ВОЛГОГРАД | Тел.: (8442) 260-048 • volgograd@prosoft.ru • www.prosoft.ru |
| ЕКАТЕРИНБУРГ | Тел.: (343) 376-2820; 356-5111 • Факс: (343) 310-0106 • info@prosoftsystems.ru • www.prosoftsystems.ru |
| КАЗАНЬ | Тел.: (843) 203-6020 • Факс: (843) 203-6020 • info@kzn.prosoft.ru • www.prosoft.ru |
| КРАСНОДАР | Тел.: (861) 224-9513 • Факс: (861) 224-9513 • krasnodar@prosoft.ru • www.prosoft.ru |
| Н. НОВГОРОД | Тел.: (831) 215-4084 • nnovgorod@prosoft.ru • www.prosoft.ru |
| НОВОСИБИРСК | Тел.: (383) 202-0960; 335-7001/7002 • Факс: (383) 230-2729 • info@nsk.prosoft.ru • www.prosoft.ru |
| ОМСК | Тел.: (3812) 286-521 • Факс: (3812) 315-294 • omsk@prosoft.ru • www.prosoft.ru |
| ПЕНЗА | Тел.: (8412) 494-971 • Факс: (8412) 494-971 • penza@prosoft.ru • www.prosoft.ru |
| САМАРА | Тел.: (846) 277-9166 • Факс: (846) 277-9165 • info@samara.prosoft.ru • www.prosoft.ru |
| УФА | Тел.: (347) 292-5216/5217 • Факс: (347) 292-5218 • info@ufa.prosoft.ru • www.prosoft.ru |
| ЧЕЛЯБИНСК | Тел.: (351) 239-9360 • chelyabinsk@prosoft.ru • www.prosoft.ru |

PROSOFT®



Рис. 4. Aurora Tigon

v2 – 1,2 ГГц, тогда как максимально возможная тактовая частота отдельных ядер процессора может достигать пикового значения 3,4 ГГц.

Для нашей задачи были включены только три вычислительных лезвия. Средствами службы Hyper-V на серверах задействованы виртуальные машины с конфигурацией, позволившей разместить все компоненты виртуальной структуры испытательного комплекса. Для каждой виртуальной машины было выделено по 8 процессоров и 16 Гбайт оперативной памяти. В качестве базовой операционной системы виртуальных машин использовалась Microsoft Windows Server 2012 R2. Организована виртуальная сеть, чтобы все виртуальные машины могли иметь взаимный сетевой доступ. Средствами операционной системы на отдельном сервере организовано общее дисковое пространство для хранения обрабатываемых данных.

В зависимости от нагрузок и размеров обрабатываемых данных указанное решение может быть масштабировано до мощного вычислительного центра с производительностью до 1020 Тфлопс в одной стойке (рис. 4). Водяное охлаждение позволяет построить максимально энергоэффективную систему с возможностью отказа от чиллеров и перехода на технологию Free Cooling, а также повторного использования тепловой энергии.

В качестве основного программного продукта для обработки информации АСУ ТП использовалось ПО компании ICONICS.

ПРОГРАММНОЕ РЕШЕНИЕ ДЛЯ ЗАДАЧ BIG DATA

Применительно к промышленному производству и типовым задачам, возникающим при работе с большими данными, на индустриальных объектах целесообразно использовать программные решения, которые являются развитием современных систем контроля и визуализации технологических процессов. Такой вывод обусловлен тем, что многие компоненты системы изначально адаптированы к типам данных, которые циркулируют в системах автоматизации промышленных предприятий. В каждой промышленной сфере, будь то энергетика, добыча и переработка нефти и газа, переработка сырья или серийное производство изделий, есть индивидуальные особенности. Чаще всего они связаны со спецификой технологических процессов, которые необходимо автоматизировать. Но практически все автоматизированные системы построены на базе одной архитектуры и имеют единые задачи автоматизации. В связи с этим в большинстве случаев могут применяться программные продукты одного типа. Одним из них является SCADA-система GENESIS64 (рис. 5) и сервер исторических данных Hyper Historian от мирового лидера по разработке программного обеспечения для автоматизации – фирмы ICONICS. Специализированный продукт Hyper Historian обладает набором функций, которые покрывают все потребности в решении задач обработки технологических данных.

В архитектуре Hyper Historian можно выделить следующие компоненты:

- Универсальный коллектор данных обеспечивает работу с информацией разных типов и обмен по стандартным протоколам OPC, SNMP, BACnet, возможно подключение к базам данных и формирование запросов. Здесь нужно отметить средства использования различных фильтров данных, таких как выделение макси-

муму, минимума, среднего значения, скользящего среднего и т.д.

- Средства подключения к облачным данным на основе публичных или частных облаков, использующих технологию Microsoft Azure™. Уже многие устройства способны передавать информацию в облако, что актуально для распределённых систем сбора данных и подвижных объектов.
- Модуль конфигурирования расчётных задач – обширная библиотека функций вычисления и преобразования данных. Пользователь может сам сконфигурировать сложный расчёт, исполнение которого вызывается периодически или по определённому условию. Операции можно выполнять как над числовыми значениями, так и над строковыми переменными, значениями времени и даты. Это хороший инструмент для разработки решений прикладных задач, специфичных для той или иной отрасли.
- Система архивирования данных реального времени с оптимизацией процесса записи событий в буферную оперативную память для получения большей производительности при пиковых нагрузках. Помимо этого реализовано стандартное логирование и запись данных в файл на жёсткий диск.
- Механизм резервирования процессов приёма и накопления данных – для обеспечения целостности данных за счёт избыточности программно-аппаратной реализации.
- Интерфейс взаимодействия с клиентскими запросами: оперативный контроль, подготовка отчётов, аналитические расчёты, взаимодействие с другими системами через открытый интерфейс.

Средствами операционной системы на физических серверах были задействованы виртуальные машины (Hyper-V) заданной конфигурации. Для обеспечения сетевого взаимодействия программных компонентов системы была органи-

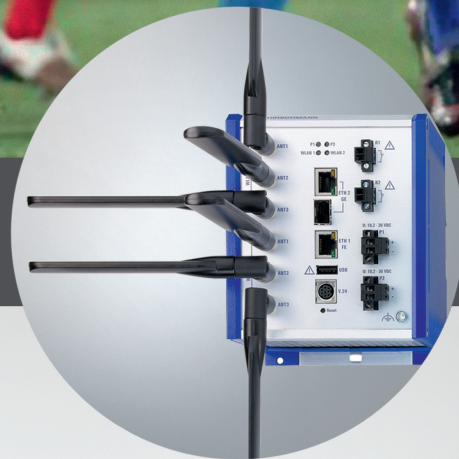


Рис. 5. Отображение данных в Hyper Historian



WLAN проходит без помех

Clear Space® — запатентованная технология получения чистого сигнала в шумных средах



Серия Hirschmann OpenBAT

Беспроводное оборудование стандарта IEEE 802.11n (Wi-Fi)

- 1 или 2 радиомодуля IEEE 802.11a/b/g/h/n
- Скорость передачи до 450 Мбит/с
- Технологии MIMO 3x3, MESH, WDS
- -40...+75°C, конформное покрытие
- Внутреннее и внешнее исполнение IP40/IP67

Вся необходимая инфраструктура:

BAT-C – простой и компактный клиент сети

Антенны, кабели, грозозащита

BAT-Controller – аппаратный централизованный контроллер точек доступа

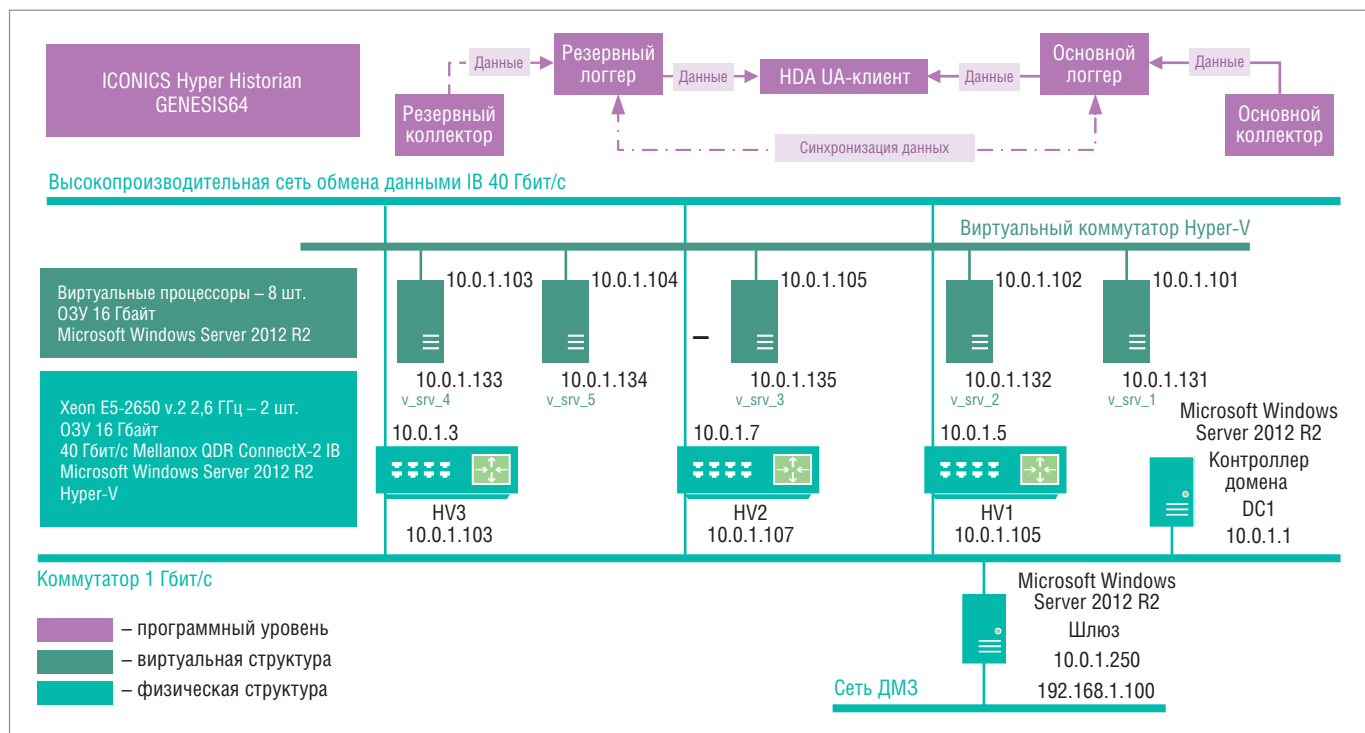
BAT-Planner – ПО для расчёта зон покрытия и скоростей передачи на плане объекта



PROSOFT®

ОФИЦИАЛЬНЫЙ ДИСТРИБЬЮТОР ПРОДУКЦИИ HIRSCHMANN

- | | |
|---------------------|--|
| МОСКВА | Тел.: (495) 234-0636 • Факс: (495) 234-0640 • info@prosoft.ru • www.prosoft.ru |
| С.-ПЕТЕРБУРГ | Тел.: (812) 448-0444 • Факс: (812) 448-0339 • info@spb.prosoft.ru • www.prosoft.ru |
| АЛМА-АТА | Тел.: (727) 220-7140/7141 • sales@kz.prosoft.ru • www.prosoft-kz.com |
| ВОЛГОГРАД | Тел.: (8442) 260-048 • volgograd@prosoft.ru • www.prosoft.ru |
| ЕКАТЕРИНБУРГ | Тел.: (343) 376-2820; 356-5111 • Факс: (343) 310-0106 • info@prosoftsystems.ru • www.prosoftsystems.ru |
| КАЗАНЬ | Тел.: (843) 203-6020 • Факс: (843) 203-6020 • info@kzn.prosoft.ru • www.prosoft.ru |
| КРАСНОДАР | Тел.: (861) 224-9513 • Факс: (861) 224-9513 • krasnodar@prosoft.ru • www.prosoft.ru |
| Н. НОВГОРОД | Тел.: (831) 215-4084 • nnovgorod@prosoft.ru • www.prosoft.ru |
| НОВОСИБИРСК | Тел.: (383) 202-0960; 335-7001/7002 • Факс: (383) 230-2729 • info@nsk.prosoft.ru • www.prosoft.ru |
| ОМСК | Тел.: (3812) 286-521 • Факс: (3812) 315-294 • omsk@prosoft.ru • www.prosoft.ru |
| ПЕНЗА | Тел.: (8412) 494-971 • Факс: (8412) 494-971 • penza@prosoft.ru • www.prosoft.ru |
| САМАРА | Тел.: (846) 277-9166 • Факс: (846) 277-9165 • info@samara.prosoft.ru • www.prosoft.ru |
| УФА | Тел.: (347) 292-5216/5217 • Факс: (347) 292-5218 • info@ufa.prosoft.ru • www.prosoft.ru |
| ЧЕЛЯБИНСК | Тел.: (351) 239-9360 • chelyabinsk@prosoft.ru • www.prosoft.ru |



Условные обозначения: HV1...HV3 – физические серверы; v_srv_1... v_srv_5 – виртуальные серверы; HDA UA-клиент – приложение для доступа к историческим данным с унифицированной архитектурой; ДМЗ – демилитаризованная зона.

Рис. 6. Схема тестового стенда

зована виртуальная сетевая структура. Для каждого виртуального сервера была определена логическая роль в системе обработки и хранения данных.

ТЕСТИРОВАНИЕ

Для проведения теста производительности (рис. 6) был выбран ряд типовых задач. Учитывая то, что тестовая система строилась на кластерной основе с использованием виртуальных машин, все итоговые нагрузочные табличные значения стоит рассматривать в разрезе нагрузок на виртуальный многоядерный процессор. Преимуществом такого подхода является оптимальная балансировка необходимых ресурсов для выполнения конкретных задач.

Оценка загрузки системы при изменении дискретности значений в источнике данных

Описание сценария:

- на имитаторе данных создаётся OPC-сервер с 10 000 аналоговых параметров;
- на стороне приёмника данных создаётся 10 000 тегов, настраивается OPC DA-интерфейс на сбор данных из источника;
- инициируется генерация значений всех параметров источника данных с дискретностью : 5, 1, 0,5, 0,2 секунды;
- на стороне приёмника данных оценивается полнота архива и загрузка системы при разной дискретности приёма: уровень загрузки процессора и

свободной оперативной памяти, скорость увеличения архива за 10 минут;

- на стороне приёмника данных формируются запросы с выборкой 10 тыс., 100 тыс., 1 млн значений к БД архива и замеряется время их выполнения (табл. 1); запросы должны ограничивать выборку по глубине таким образом, чтобы размер архива не влиял на длительность выполнения запроса (например, ограничение по временному периоду, количеству параметров и т.д.).

Оценка загрузки системы при увеличении количества собираемых параметров

Описание сценария:

- на имитаторе данных создаётся OPC-сервер с 10 000 аналоговых параметров;
- инициируется генерация значений с дискретностью 1 с;
- на стороне приёмника данных настраивается и запускается сбор данных по OPC DA из источника с количеством тегов: 10 000, 30 000, 50 000, 100 000;
- на стороне приёмника данных оценивается полнота архива и загрузка системы при разном количестве тегов: уровень загрузки процессора и свободной оперативной памяти, скорость увеличения архива за 10 минут;

Оценка загрузки системы – OPC DA-сервер

Таблица 1

| Параметры | Без нагрузки | Период генерации значений, с | | | |
|---|--------------|------------------------------|---------|---------|---------|
| | | 5 | 1 | 0,5 | 0,2 |
| Загрузка процессора при сохранении данных, % | – | 2,40 | 2,90 | 3,30 | 4,50 |
| Использование оперативной памяти при сохранении данных, Мбайт | – | 910,00 | 1020,00 | 1140,00 | 1280,00 |
| Загрузка процессора при сборе данных, % | – | 0,50 | 1,10 | 2,70 | 6,00 |
| Использование оперативной памяти при сборе данных, Мбайт | – | 460,00 | 620,00 | 710,00 | 950,00 |
| Изменение размера использования HDD за 10 мин, Мбайт | – | 40,00 | 80,00 | 160,00 | 360,00 |
| Время обработки 10 000 значений (TrendWorX Viewer), с | 1,50 | 1,50 | 1,50 | 1,50 | 1,50 |
| Время обработки 100 000 значений (TrendWorX Viewer), с | 7,40 | 7,40 | 7,40 | 7,40 | 7,40 |
| Время обработки 1 000 000 значений (TrendWorX Viewer), с | 63,00 | 63,00 | 63,00 | 63,00 | 63,00 |



Встраиваемые решения MEN

Защищённые компьютерные платы и системы для работы
в жёстких условиях эксплуатации и для ответственных применений

Высокое качество продукции в соответствии с ISO 9001/14001, AN/AS 9100, IRIS

Высокая надёжность в соответствии с EN 50155, DO-254, E1

Обеспечение уровней безопасности до SIL 4, DAL-A

Компьютерные модули Rugged COM Express® (VITA 59) и ESMexpress®

Платы в форматах CompactPCI®/PlusIO/Serial и VME

Мезонинные модули PMC, XMC, M-Module™ I/O

Защищённые коммутаторы Ethernet

Встраиваемые и панельные компьютеры



ОФИЦИАЛЬНЫЙ ДИСТРИБЬЮТОР ПРОДУКЦИИ MEN MIKRO ELEKTRONIK

| | |
|---------------------|---|
| МОСКВА | Тел.: (495) 234-0636 • Факс: (495) 234-0640 • info@prosoft.ru • www.prosoft.ru |
| С.-ПЕТЕРБУРГ | Тел.: (812) 448-0444 • Факс: (812) 448-0339 • info@spb.prosoft.ru • www.prosoft.ru |
| АЛМА-АТА | Тел.: (727) 220-7140/7141 • sales@kz.prosoft.ru • www.prosoft-kz.com |
| ВОЛГОГРАД | Тел.: (8442) 260-048 • volgograd@prosoft.ru • www.prosoft.ru |
| ЕКАТЕРИНБУРГ | Тел.: (343) 376-2820; 356-5111 • Факс: (343) 310-0106 • info@prosoftsystems.ru • www.prosoftsystems.ru |
| КАЗАНЬ | Тел.: (843) 203-6020 • Факс: (843) 203-6020 • info@kzn.prosoft.ru • www.prosoft.ru |
| КРАСНОДАР | Тел.: (861) 224-9513 • Факс: (861) 224-9513 • krasnodar@prosoft.ru • www.prosoft.ru |
| Н. НОВГОРОД | Тел.: (831) 215-4084 • novgorod@prosoft.ru • www.prosoft.ru |
| НОВОСИБИРСК | Тел.: (383) 202-0960; 335-7001/7002 • Факс: (383) 230-2729 • info@nsk.prosoft.ru • www.prosoft.ru |
| ОМСК | Тел.: (3812) 286-521 • Факс: (3812) 315-294 • omsk@prosoft.ru • www.prosoft.ru |
| ПЕНЗА | Тел.: (8412) 494-971 • Факс: (8412) 494-971 • penza@prosoft.ru • www.prosoft.ru |
| САМАРА | Тел.: (846) 277-9166 • Факс: (846) 277-9165 • info@samara.prosoft.ru • www.prosoft.ru |
| УФА | Тел.: (347) 292-5216/5217 • Факс: (347) 292-5218 • info@ufa.prosoft.ru • www.prosoft.ru |
| ЧЕЛЯБИНСК | Тел.: (351) 239-9360 • chelyabinsk@prosoft.ru • www.prosoft.ru |

Оценка загрузки системы при изменении количества собираемых параметров

Таблица 2

| Параметры | Количество тегов | | | |
|---|------------------|---------|---------|---------|
| | 10 000 | 30 000 | 50 000 | 100 000 |
| Загрузка процессора при сохранении данных, % | 3,60 | 9,60 | 14,50 | 22,80 |
| Использование оперативной памяти при сохранении данных, Мбайт | 1240,00 | 2050,00 | 2520,00 | 3340,00 |
| Загрузка процессора при сборе данных, % | 4,60 | 8,60 | 11,90 | 19,50 |
| Использование оперативной памяти при сборе данных, Мбайт | 930,00 | 1420,00 | 1790,00 | 2250,00 |
| Время обработки 100 000 значений, с | 7,60 | 7,70 | 7,80 | 7,60 |
| Время обработки 1 000 000 значений, с | 63,00 | 62,00 | 65,00 | 61,00 |

Оценка скорости чтения параметров

Таблица 3

| Тест на количество читаемых значений | | | | | |
|--|--------------|-------------------|--------------------|--------------------|------------------------|
| Количество значений | Менее 50 000 | 50 000... 150 000 | 150 000... 450 000 | 450 000... 350 000 | 1 350 000... 4 050 000 |
| Время обработки, с | 3,40 | 9,30 | 29,20 | 86,80 | 254,40 |
| Тест на количество запросов | | | | | |
| Количество запросов с выборкой более 500 000 | 3 | 6 | 10 | 30 | |
| Максимальное время выполнения, с | 37 | 68 | 107 | 384 | |

Влияние выполнения расчётов на загрузку системы

Таблица 4

| Параметр | С расчётом | Без расчёта |
|---|------------|-------------|
| Загрузка процессора, % | 65,70 | 7,10 |
| Использование оперативной памяти, Мбайт | 7200 | 1870 |

● на стороне приёмника данных формируются запросы с выборкой 100 тыс., 1 млн значений к БД архива и замеряется время их выполнения; запросы должны ограничивать выборку (табл. 2) по глубине таким образом, чтобы размер архива не влиял на длительность выполнения запроса (например, ограничение по временному периоду, количеству параметров и т.д.).

Оценка скорости чтения параметров

Описание сценария:

- на стороне приёмника данных запускается запрос на выборку значений из архива (не менее) 50 000, 150 000, 450 000, 1 350 000, 4 050 000 значений;
- оценивается время выполнения каждого из запросов;
- на стороне приёмника данных параллельно запускается несколько запросов на выборку не менее 500 000 значений из архива: 3, 6, 10, 30 запросов (при наличии возможности);
- оценивается время выполнения запросов (наибольшее значение); запросы должны ограничивать выборку (табл. 3) по глубине таким образом, чтобы размер архива не влиял на длительность выполнения запроса (например, ограничение по временному периоду, количеству параметров и т.д.).

Оценка выполнения расчётов и их влияние на загрузку системы

Описание сценария:

- на имитаторе данных создаётся ОРС-сервер с 10 000 аналоговых параметров;
- создаётся скрипт, позволяющий генерировать значения параметров;
- на стороне приёмника данных подготавливается 10 000 тегов с подпиской на параметры источника данных;
- на стороне приёмника данных подготавливается 5 000 расчётных тегов, рассчитываемых по формуле: значение расчётного тега = среднее значение [исходный тег ОРС];
- для расчёта используется массив из 500 последних архивных значений исходного тега ОРС;
- расчёт должен производиться в режиме «по изменению»;
- на стороне приёмника данных подготавливается 5 000 расчётных тегов, рассчитываемых по формуле: значение расчётного тега = среднеквадратичное отклонение [исходный тег ОРС];
- для расчёта используется массив из 500 последних архивных значений исходного тега ОРС;
- расчёт должен производиться в режиме «по изменению»;
- инициируется генерация значений с дискретностью 1 секунда длительностью 10 минут;

● на стороне приёмников данных оценивается полнота архива сборных и расчётных параметров и загрузка системы: уровень загрузки процессора, свободной оперативной памяти (табл. 4);

● производится сравнение параметров загрузки системы с аналогичным сценарием без выполнения расчётов.

Как видно из приведённых примеров, система успешно справлялась со всеми запросами и изменяемыми нагрузками. Хорошим знаком является то, что с увеличением количества обрабатываемой информации в относительных цифрах происходит оптимизация работы кластера, то есть работа под нагрузкой и с большими данными для кластера является оптимальным режимом функционирования.

РЕКОМЕНДАЦИИ

Выбор аппаратной части следует осуществлять с привязкой к размерности предполагаемых задач. Для рассмотренных в ходе тестирования задач нет необходимости использовать большие массивно-параллельные платформы для вычислений, достаточно двух либо трёх двухsocketных вычислительных узлов.

Для эффективного решения задач большей размерности (> 500 тыс. точек) либо решения оптимизационных задач целесообразно использование вычислительной мощности нескольких узлов, блейд-системы отлично подходят для реализации проектов в области АСУ ТП, когда дело касается обработки больших данных. Также значительным преимуществом такого «железа» становится отказоустойчивость и возможность линейного наращивания вычислительных способностей.

Стоит отметить, что выпущена новая версия 10.95 программного обеспечения от ICONICS, которая уже показала значительное повышение производительности по результатам вычислений, но на момент написания статьи она находилась на бета-тестировании. Официальные данные по производительности текущего релиза можно найти на DVD-диске компании ПРОСОФТ или ICONICS соответствующей версии. Документ называется “Hyper Historian Performance”.

Всё, что касается работы программного комплекса, будет раскрыто в следующей статье по этой тематике. ●

Авторы – сотрудники фирмы ПРОСОФТ
Телефон: (495) 234-0636
E-mail: info@prosoft.ru

Анализ технологических показателей в реальном времени

Решения на базе программных продуктов ICONICS



ЧТО?

- Управление эксплуатацией оборудования
- Снижение затрат
- Энергоменеджмент

КАК?

- Диагностика состояния оборудования с возможностью прогнозирования сбоев. Учёт наработки, экспертные карты, вероятностный анализ
- Анализ нештатных режимов. Частота возникновения, поиск взаимосвязи, анализ времени реакции персонала
- Анализ потребления энергоресурсов. Данные о потреблении в реальном времени, сравнение с идеальной моделью и плановыми показателями, сравнение с историческими данными, индикаторы энергоэффективности. Поддержка анализа в рамках энергоменеджмента по ГОСТ 50001:2011



Make the Invisible Visible™

PortalWorX Productivity Analytics Facility AnalytiX Energy AnalytiX Alarm Analytics



Тел.: +7 (495) 232-1817
Факс: +7 (495) 232-1649
Эл. почта: info@norvix.ru

Официальный партнёр
компании ПРОСОФТ
www.norvix.ru