

Оценка производительности рабочих станций

Пособие для начинающих

© 2002 CAD/CAM Publishing, Inc. COMPUTER AIDED DESIGN REPORT™ (No. 1, 2002)

Мы часто слышим, что цены на компьютеры падают, но это только одна сторона дела. Затраты на компьютерное оборудование по-прежнему составляют значительную часть бюджета на САПР. В отличие от мобильных телефонов, рабочие станции пока еще не раздают бесплатно. Хотя покупатели теперь получают больше производительности на каждый доллар, вложенный в аппаратное обеспечение, расходы по-прежнему велики. Поэтому только немногие счастливицы могут позволить себе роскошь покупать все самое лучшее, в то время как большая часть тех, кто приобретает компьютеры для инженерных приложений, вынуждена идти на компромиссы.

Сделать хороший выбор, не имея надежных данных, чрезвычайно трудно. Чтобы решить эту проблему, многие администраторы CAD-систем руководствуются результатами тестирования производительности с помощью специальных программ, которые измеряют быстродействие сравниваемых компьютеров и их подсистем при выполнении различных задач.

Правильно разработанные и грамотно используемые тесты могут дать ценную информацию для сравнения и конфигурирования новых компьютеров. Результаты корректных тестов полезны также и при модернизации имеющегося оборудования. **Обратное тоже верно:** неправильно разработанные или использованные тесты могут привести к тому, что ваши аппаратные доллары отправятся по неверному адресу. Учитывая высокие ставки, мы предлагаем это пособие по тестированию, созданное на основе собственного опыта.

Так ли необходимо тестирование?

Когда речь заходит о производительности, скептики иногда утверждают, что благодаря революции, произведенной *Windows*, весь вопрос выбора оборудования теперь сводится только к приобретению максимально быстрой системы на базе процессора *Pentium*. По их словам, оценка производительности сейчас не стоит затраченных усилий. Эта точка зрения представляется нам упрощенной и далекой от реальности.

Даже из арсенала рабочих станций на базе *Intel* покупатели должны выбрать линейку процессоров: 32-разрядные *Xeon* и *Pentium 4* или новый 64-битный *Itanium*. Старый добрый *Pentium III* тоже все еще доступен, особенно в новом классе портативных рабочих станций. Еще более усложняет картину *AMD* – грозный конкурент *Intel* в нише 32-разрядных домашних компьютеров. И *AMD*, и *Intel*

предлагают чипы для рабочих станций, работающие с различной тактовой частотой, причем как в однопроцессорном, так и двухпроцессорном варианте. Кроме того, предлагается и память различной архитектуры, включая *RAMBUS DRAM* и *DDR* (с двойной частотой).

В области САПР для интегральных схем выбор сейчас стал даже еще более сложным. Здесь пока доминируют системы на базе *Unix*, эффективность которых в значительной степени определяется архитектурой процессора. Покупатель рабочей станции должен решить, действительно ли более дорогие *Unix*-системы обеспечат лучшую отдачу вложенных денег либо стоит повернуться лицом к *Linux* на базе *Intel* или *AMD*.

Наконец, есть принципиальные моменты, определяющие конфигурацию. Углубленные тесты показывают, что наличие достаточного объема оперативной памяти более важно для производительности, чем частота процессора. Но какое именно количество мегабайтов можно считать достаточным объемом? Помимо этого, необходимо выбрать и другие компоненты. Вы должны решить, сколько потратить на графическую карту, дисководы, другие внешние устройства.

Лучший ответ на все эти вопросы дает анализ результатов хороших тестов. Правильно проведенное тестирование позволяет исключить множество рисков в процессе принятия решения о приобретении компьютеров. Научный подход может дать экономию в сотни долларов на каждую рабочую станцию. Для крупных организаций, где одновременно приобретается большое количество компьютерной техники, грамотное тестирование способно сэкономить тысячи долларов.

Осторожно: бесплатные данные!

В журналах или в Сети можно найти целые тома различных рейтингов несметного числа производителей компьютеров. Эту информацию предоставляют разработчики аппаратуры, крупные продавцы компьютеров, компьютерные издания (включая *Ziff-Davis* в мире ПК и *ConnectPress* для рабочих станций), агентства типа *Standard Performance Evaluation Corporation (SPEC)*. Но вы получаете то, за что платите, – к этому изобилию стоит относиться с осторожностью.

Обычная проблема множества опубликованных рейтингов заключается в том, что у разработчиков “железа” слишком много возможностей влиять на условия проведения испытаний. В случае *SPEC* и ее многочисленных подразделений, продавцы оборудования сами проводят испытания и предоставляют результаты. Что касается тестов, проводимых журналами, –

не секрет, что поставщики часто сами конфигурируют тестируемые системы. В любом случае, именно продавцы решают, какое оборудование предоставить, что зачастую позволяет сравнивать только яблоки с виноградом. Есть множество примеров того, как сравниваемые системы комплектовались различными ключевыми компонентами, включая разный объем памяти, графические карты или даже процессоры с разной тактовой частотой. По таким отчетам крайне трудно выявить, почему один компьютер работает лучше другого.

Как мы уже сказали, программы, используемые в популярных тестах, могут быть полезными и для покупателя рабочей станции. Многие из этих программ можно скачать из Интернета или заказать на CD с доставкой по почте (подробнее о публично доступных тестах – в конце статьи).

Покупателям, которые полагаются на опубликованные данные, тоже следует ознакомиться с тестовыми программами, желательно путем их использования. Это лучший способ понять, производительность чего именно в действительности измеряют те или иные тесты, какими моделями манипулируют, какие процедуры и программные модули задействуют и как механизмы подсчета могут влиять на результат.

К примеру, мы регулярно дополняем тестовые задачи, разработанные по нашему заказу, тестами приложений *Pro/ENGINEER* от *SPEC's Application Performance Characterization Group (SPECapc)*. Мы пришли к выводу, что их тесты прекрасно показывают, какая графическая карта будет лучшей для *Pro/E*. Однако они не столь хороши для оценки процессора и памяти – компонентов, наиболее важных для производительности машиностроительных CAD-систем, таких как *Pro/E*.

Не нравятся нам у *SPECapc* и способы подведения итогов, которые представляют собой средние значения результатов многократных подтестов, пронормированные относительно производительности старого компьютера, называемого системой координат. С нашей точки зрения, более предпочтительны результаты, выраженные в единицах времени, например секундах, что позволяет лучше почувствовать связь каждого теста с реальной производительностью. Поэтому после выполнения теста *Pro/E* мы исключаем средние значения и приводим фактическое время, необходимое для выполнения отдельных тестовых задач. Такие данные доступны только при самостоятельном проведении тестирования.

Результаты испытаний *SPECapc* используются нами прежде всего для сравнения производительности графических карт. Продавцы компьютеров часто представляют эти тесты как характеристику общей производительности *Pro/E*, что, как минимум, вводит пользователя в заблуждение.

Лучший способ тестирования

Ключевое правило тестирования: лучший тест – тот, который вы разработали сами. Главное преимущество тестовых программ собственного производства состоит

в том, что вы точно знаете, что конкретно они будут делать, как подсчитываются результаты и каковы условия, в которых эти результаты получены. Собственные тесты могут быть построены таким образом, чтобы моделировать рабочие нагрузки, характерные именно для ваших задач. Поэтому результаты тестов, разработанных специально под ваши задачи, дадут наиболее надежные данные для оценки как различных классов, марок или моделей рабочих станций в целом, так и их отдельных компонентов (памяти, графики и пр.).

Второе преимущество собственных тестов – их гибкость. Если вы создаете тест сами, у вас есть и возможность улучшать его. Он будет развиваться и удовлетворять ваши потребности, даже если эти потребности будут меняться.

Конечно, у самодельных тестов есть существенный недостаток – их надо создавать. На это нужно тратить ресурсы, в том числе и деньги, если для этого приходится нанимать разработчика. Но, ввиду того, что надежные данные тестирования могут помочь сохранить тысячи долларов, выгода может перевесить расходы. Ну а у тех, кто не хочет обременять себя, нет другого выхода, кроме как полагаться на тесты, разработанные кем-то.

Разработка собственных тестов

До начала разработки собственных тестов необходимо выявить специфику рабочей нагрузки вашего оборудования и сформировать обширный список задач-кандидатов. Начните с основных приложений. В пределах каждого приложения следует выявить используемые модули, а также примерно определить, насколько часто они используются. Например, если ваши инженеры, работающие с *Pro/E*, обычно запускают *Pro/Process for Assemblies* или *Behavioral Modeling Extensions*, то и в вашем тесте они должны присутствовать. Если регулярно применяются *PhotoWorks* или *Animator*, ваш тест *SolidWorks* обязательно должен включать операции, задействующие эти программы.

Необходимо также обратить внимание на типы моделей. Подберите ряд подходящих моделей, в наибольшей степени соответствующих по размеру и сложности текущим или планируемым нагрузкам. Работают ли ваши инженеры преимущественно с отдельными деталями, со сборками или же и с тем и с другим? Сколько вложенных подборок содержит типичное изделие? Имеют ли ваши модели поверхности произвольной формы или состоят, главным образом, из плоских граней?

Строят ли ваши инженеры сечения твердотельных моделей (графический акселератор может увеличить производительность этой функции)? Какой тип рендеринга они предпочитают? Если они работают с моделью в режиме сокрытия невидимых линий, то этот режим должен быть включен в тест. Используются ли текстуры



Тест Ugs-01 из набора SPECviewpref 7.0

только для того, чтобы модели выглядели круче, или это действительно необходимо для показа проекта заказчику?

Как только вы провели эту инвентаризацию рабочей нагрузки, сфокусируйте процедуры испытаний на наиболее важных вещах. Тесты, которые включают все возможные CAD- или инженерные функции, занимают слишком много времени, да и для интерпретации результатов потребуются докторская степень в области статистики. Для сужения этого списка необходимо идентифицировать параметры, критичные для производительности на ваших задачах. Если система тормозит на чертежах, больших сборках, моделях сложных криволинейных форм или каких-то процедурах рендеринга, – удостоверьтесь, что ваши тесты включают в себя такие операции. Нет смысла включать в тестирование процедуры, которые выполняются быстро и не требуют больших ресурсов.

Продолжительность выполнения тестов должна быть достаточной для точного отражения производительности, но не слишком большой. Мы полагаем, что длительность теста более 15 минут чрезмерна, но это зависит от вашей специфики.

После определения рабочей нагрузки необходимо рассмотреть другие требования. Хорошие тесты удобны. Наиболее часто результатом тестирования будет количество времени, необходимое компьютеру для выполнения определенных операций. Однако сидеть перед компьютером с секундомером вряд ли рационально. Поэтому тестовая программа должна включать подпрограмму для считывания системного времени и вычисления времени выполнения как теста в целом, так и его отдельных этапов. Надо позаботиться и о том, чтобы не вводить результаты испытаний в таблицу вручную, – удостоверьтесь, что эти данные автоматически записываются в файл для последующего анализа.

Формируйте результаты тестов в удобной для понимания форме. Не устанавливайте в тестовой программе запись времени начала и конца выполнения

процедур – это потребует затем дополнительных вычислений для определения длительности операций. Вместо этого заставьте саму программу делать вычитание и записывать длительность выполнения.

Обратите внимание и на формат представления данных. Самый простой путь – указывать время в секундах, но при желании можно отражать время выполнения в часах, минутах и секундах или даже в долях секунды. Также стоит обдумать, хотите ли вы расположить данные по столбцам или по строкам, в виде электронной таблицы или в обычном текстовом файле. Вывод данных непосредственно в формате электронной таблицы позволит сэкономить пару шагов. Сохранение данных в текстовом файле может оказаться проще в реализации.

Сами приложения, которые вы используете для тестирования производительности, в значительной степени могут помочь в написании тестов. Так, *CATIA V5* и *SolidWorks* могут автоматически формировать сценарии на языке *Visual Basic* при построении модели или чертежа. Эта возможность сокращает время, необходимое для разработки тестовых программ. В отличие от этого тесты для *Solid Edge* или *Inventor* нужно составлять вручную, что требует больше труда. Система *Pro/ENGINEER* позволяет сформировать тесты с помощью *trail*-файлов, которые подобно макросам записывают все шаги построения модели и манипуляции с ней.

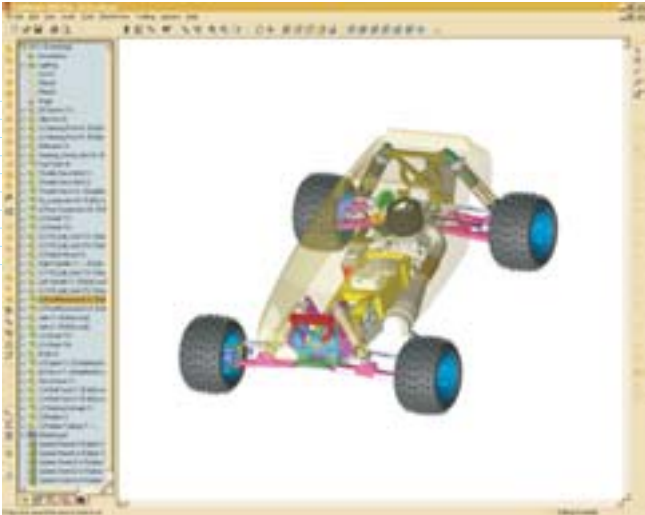
С течением времени вам потребуется адаптировать тест к новым версиям программного обеспечения. Может возникнуть необходимость перезаписать тестовый сценарий, чтобы он соответствовал обновленному ПО.

Некоторые соображения о процессе тестирования

Самое важное при проведении тестов – обеспечить максимальную достоверность данных. Убедитесь, что



Так выглядит тот самый копир из популярного теста производительности Pro/ENGINEER



Один из этапов прикладного теста SPECarc for SolidWorks 2001Plus

тестируемые системы сконфигурированы одинаковым образом. Это нетрудно, если тестируется какой-то один компонент – например, подбирается самая быстрая видеокарта для данного приложения на данном типе компьютера. Все, что необходимо сделать в этом случае, – выбрать тестовую систему и выполнять все тесты на ней, заменяя только сравниваемые компоненты.

Самодельные тесты должны выполняться несколько раз за каждый сеанс, причем результаты каждого прохода нужно записывать отдельно. Эти результаты могут иногда различаться. Асы тестирования очень не любят подобные различия и стремятся к единственно верному набору цифр. Поэтому, например, тесты SPECarc повторяют все процедуры по три раза и затем вычисляют среднее значение.

Однако усреднение результатов может скрыть ценную информацию. Например, вполне вероятно, что производительность системы на втором и третьем проходах теста в пределах одного сеанса будет выше за счет сохранения данных в кэш-памяти.

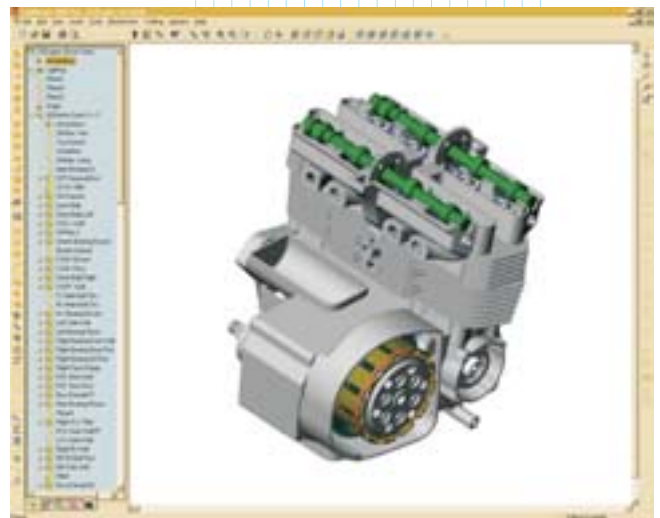
Используйте как минимум два разных теста, даже если для этого придется дополнить ваши собственные тесты общедоступными. Если по результатам нескольких тестов производительность вызывает разочарование, это может указывать на проблемы с ОС или с конфигурацией системы. Не забывайте перезагружать систему между испытаниями. Сохранение единого порядка выполнения тестов повышает надежность получаемых результатов. Решите для себя, будете ли вы перезагружать приложение после каждого прохода одного и того же теста. Мы также всегда фиксируем порядок, в котором мы выполняли тесты, число проходов каждого теста, перезагружался ли приложение и когда перезагружался компьютер.

Если вы собираетесь сравнивать результаты испытаний нескольких систем, старайтесь использовать максимально близкие конфигурации: одинаковый объем

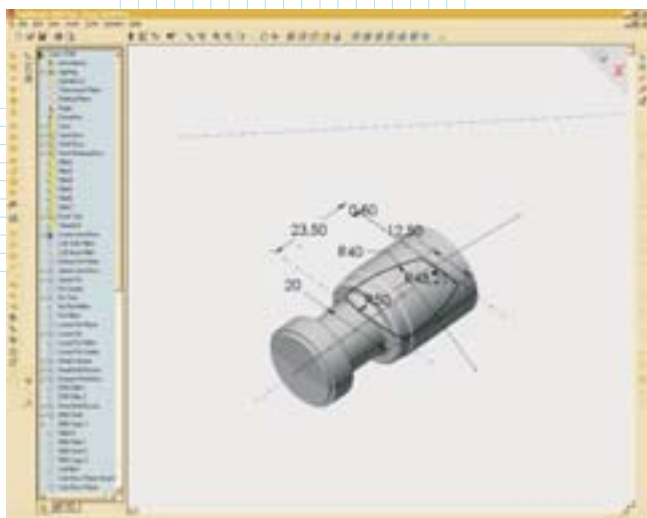
памяти, такую же (или, по крайней мере, подобную) графическую карту. Везде, где это возможно, следует использовать одну и ту же версию операционной системы. Если при тестировании происходит частое обращение к жесткому диску, то характеристики дисков должны быть сходными, особенно объем диска, частота вращения пластин, размер дискового кэша и свободного пространства.

Использование в тестах разных версий программного обеспечения может исказить результаты. Удостоверьтесь, что используются одинаковые версии программ, включая идентичные *Build number* или *Date codes*, патчи и сервис-паки. Поскольку разработчики непрерывно модернизируют программы, они добавляют коды, которые иногда улучшают или ухудшают производительность. Настройки, сделанные в приложении, также могут влиять на производительность, поэтому они должны быть приведены к единому образцу после каждой переинсталляции.

Систематизируйте результаты испытаний в таблице или базе данных сразу после того, как вы их получили. Это позволит избежать непоследовательности и поможет быстро обнаружить аномалии в результатах, вызванные неправильными установками или чем-либо еще. Но имейте в виду, что часто именно неожиданные результаты дают наиболее ценную информацию. Например, когда мы впервые тестировали систему на базе *Pentium 4*, то были потрясены, обнаружив, что по производительности эта система с процессором *1.5GHz* не превосходила *Pentium III* с частотой *1GHz*. Сначала мы думали, что это аномалия, вызванная проблемами с конфигурацией. Но и после перезагрузки операционной системы испытания показали, что *Pentium 4* лишь немного быстрее, чем более старая модель, при гораздо более высокой цене. Владея такой информацией, компании, которые в то время приобретали компьютеры, имели возможность сберечь деньги.



Один из этапов прикладного теста SPECarc for SolidWorks 2001Plus



Один из этапов прикладного теста SPECapc for SolidWorks 2001Plus

Что может идти не так

Как любая экспериментальная работа, тестирование – сложный процесс. Полученный для тестирования компьютер может быть загружен всяким мусором, иметь сильно фрагментированное дисковое пространство, “криво” установленную операционную систему или неподходящие драйверы устройств. Когда вы тестируете систему, которая до этого уже была в эксплуатации, неплохо будет переформатировать диски и по-новой установить операционную систему. Если в процессе тестирования вы заменяете какой-либо компонент, необходимо полностью удалить его драйверы и загрузить новые.

Интерпретация результатов

Важно, чтобы при сравнении производительности разных компьютерных систем множество незначительных отличий не затмевали действительно существенного. Необходимо выявить те компоненты, которые в наибольшей степени влияют на скорость работы, а затем соотнести их стоимость с выигрышем в производительности.

Наш метод состоит в том, чтобы экстраполировать результаты тестирования на один полный день работы инженера, сравнивая стоимость времени, сэкономленного в результате модернизации аппаратного обеспечения, с ценой этой модернизации. Лучше всего показать это на примере.

Один из наших Pro/E-тестов использует сборочную модель лазерного принтера, состоящую из 848 деталей, входящих в несколько сотен подборок. В процессе тестирования Pro/E строит модель, редактирует детали, регенерирует различные компоненты, ищет пересечения деталей, вычисляет вес устройства, формирует упрощенные представления, открывает чертежи, вращает, увеличивает и уменьшает различные виды (в том числе и разрезанные), а также

производит ряд других действий. По нашим подсчетам, инженер, разрабатывающий подобное изделие, может повторять такую последовательность действий 14 раз в день.

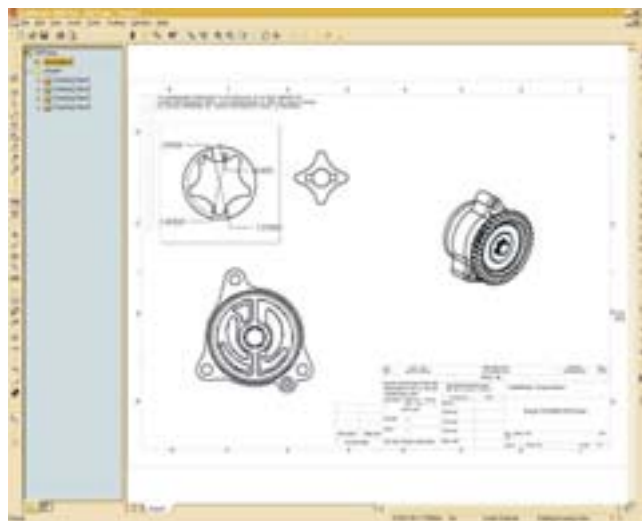
Согласно нашему первому отчету, подготовленному с использованием Pro/E-теста, рабочая станция Dell, снабженная 1Gb памяти, выполнила этот тест за 6 минут 28 секунд. Этой же системе, но с 256 Mb потребовалось почти на четыре минуты больше. Масштабируя результат на весь рабочий день путем умножения разницы на 14, получаем, что дополнительная память поможет ежедневно сберечь почти час рабочего времени инженера. На тот момент компания Dell запрашивала \$2 320 за дополнительные 768Mb (хорошая цена по тем временам, но непомерная по сегодняшним понятиям). Принимая, что издержки компании составляют 40÷65 долл. за один час работы инженера, мы приходим к выводу, что необходимо от 40 до 64 рабочих дней, чтобы окупить стоимость дополнительной памяти.

Нагрузки, которые вы выбираете для тестов, могут быть различными, отражая разные ситуации, возникающие у пользователя на протяжении рабочего дня. Такая методика позволяет как оценить оборудование, так и подготовить технико-экономическое обоснование.

Общедоступные тесты

Вполне возможно, что пользователи, планирующие создание собственных тестов, захотят вначале приобрести опыт, используя для этого общедоступные программы тестирования. Существуют два типа таких тестов: синтетические и прикладные.

Синтетические тесты – это программы, построенные по своим собственным правилам. Вы просто загружаете программу и запускаете ее. Наиболее часто такие тесты предназначены для проверки отдельных компьютерных подсистем. Например, набор тестов Viewperf, разработанный SPEC's OpenGL



Один из этапов прикладного теста SPECapc for SolidWorks 2001Plus

Graphics Performance Characterization (GPC) Group. Два из шести тестов *Viewperf – MedMCAD* и *Pro/CDRS* – дают полезную информацию о работе графической подсистемы. Вы можете бесплатно загрузить это приложение по адресу www.specbench.org/gpc/apc.static/overview.htm.

Другой достойный внимания набор синтетических тестов – *SPEC CPU2000*. Его результаты делятся на две части. Первая – результаты тестов на основе целочисленных операций (*SPECint2000*), вторая – на основе вычислений с плавающей точкой (*SPECfp2000*). Значимость этих тестов была гораздо большей во времена дебатов о *RISC*- и *CISC*-архитектуре. Тогда это был наиболее популярный инструмент межплатформенного сравнения для технических вычислений. Отрицательное свойство набора *CPU2000* состоит в том, что он стоит \$500. Если у вас есть интерес, обращайтесь по адресу www.specbench.org/cgi-bin/osgorder.

Синтетические тесты обычно не требуют каких-либо специальных лицензий, – даже при том, что они часто используют фрагменты прикладных программ, перекомпилированные для тестов. Загрузите программу – и вы в деле. Однако поведение реальных инженерных приложений может весьма отличаться от ожидаемого в соответствии с резуль-

татами такого тестирования. В идеале пользователи должны оценивать производительность компьютерных систем тестами на базе тех прикладных программ, которые они планируют использовать.

Для пользователей *MCAD* есть два главных источника прикладных тестов, которые можно загрузить из Интернета: *SPECapc* и *ConnectPress*. *SPECapc* предлагает отдельные тесты для *Pro/E*, *SolidWorks*, *Unigraphics* и *Solid Edge*. Как уже было сказано выше, эти тесты имеют тенденцию давать искаженные результаты за счет преувеличения роли графической подсистемы, но они не так уж плохи. Для получения дополнительной информации и бесплатной загрузки этих программ можно обратиться на www.specbench.org/gpc/apc.static/apcfaq.htm.

Альтернативные тесты для *SolidWorks*, *Pro/E*, а также *CATIA V5* можно получить от *ConnectPress*. Тесты *ConnectPress* представляются нам достаточно хорошими. Единственное критическое замечание – их результаты содержат так много информации, что не всегда легко разобраться, что они означают. Для получения дополнительной информации посетите сайт www.ConnectPress.com.

Успешного вам тестирования! 

Послесловие редакции

Время не стоит на месте, поэтому при подготовке этой статьи к печати наша редакция решила уточнить информацию, касающуюся публично доступных тестов. Результаты показались нам достаточно интересными, чтобы сказать о них несколько слов.

В настоящее время на смену набору тестов ***SPECviewperf 6.1.2*** пришел новый – версия ***7.0***, хотя “шестерка” тоже доступна. В версии 7.0 изменился ряд процедур тестирования, часть тестов заменена на новые. Но за прогресс приходится платить – объем инсталляционных файлов резко возрос, и вместо прежних 55Mb теперь предлагается скачать по Сети около 300Mb. Из новой версии исключены упомянутые в статье тесты *MedMCAD* и *Pro/CDRS*. Вместо них наибольший интерес для нас представляют *Ugs-01* и *Proe-01*. Обращаем ваше внимание на то, что результаты тестирования, полученные в 6-й версии, нельзя сравнивать с результатами 7-й.

Загрузить тест *SPECviewperf* можно по адресу www.specbench.org/cgi-bin/download.cgi. Однако мы столкнулись с проблемами, пытаясь закачать этот файл непосредственно из браузера. В этом случае рекомендуется обратиться к одному из зеркальных сайтов, список которых находится по адресу www.specbench.org/mirrors.html. При этом значительно удобнее и надежнее использовать *FTP*-клиент (*WS FTP*, *Windows Commander* и др), т.к. *Web*-браузер не позволяет полноценно общаться с

FTP-сервером, а также докачивать файл при обрыве связи.

Для проверки прикладных тестов в нашем распоряжении был пакет *SolidWorks 2001 Plus*. Мы загрузили тесты как с сайта ***SPEC (SPECapc for SolidWorks 2001Plus)***, так и с сайта ***ConnectPress (SolidWorks 2001 Plus Benchmark)***. При этом тест *SPECapc* доступен сразу, а для того чтобы воспользоваться услугами *ConnectPress*, необходимо заполнить анкету и подождать – в течение суток вы получите по *e-mail* пароль доступа, который действителен 48 часов с момента отправки. Как оказалось в итоге, оба теста подготовлены компанией *SolidWorks Corporation*, и нам не удалось обнаружить между ними каких-либо отличий, за исключением того, что версия от *SPEC* оказалась чуть-чуть новее.

На сайте *SPEC* прикладные тесты *Pro/E 2000i2*, *Solid Edge 11* и др. можно загрузить по адресу www.specbench.org/gpc/downloadindex.html. Там же указаны ссылки на зеркальные узлы, в том числе и *FTP*. К примеру, для нас наиболее удобным оказался итальянский <ftp://spec.uni-pv.it>, в то время как американский узел постоянно рекомендовал зайти позже.

На сайте *ConnctPress* можно найти много готовых результатов тестирования различных компьютеров. Желающих загрузить тесты сайт отправляет к странице с перечнем тестов на www.cadbench.com, которая в свою очередь отправит вас на сайты *Pro/E*, *SolidWorks* или *CATIA*. 