

СИСТЕМЫ ОБРАБОТКИ СИГНАЛОВ В РЕЖИМЕ РЕАЛЬНОГО ВРЕМЕНИ**Введение**

В радиотехнических системах (РТС) различного назначения широко используются методы цифровой обработки сигналов. По сравнению с аналоговыми методами, они обеспечивают большую гибкость в реализации, стабильность параметров при изменении условий среды, а также широкие возможности по хранению и передаче информации по каналам связи. Возможности систем цифровой обработки сигналов часто ограничиваются лишь доступной вычислительной мощностью системы обработки. В качестве вычислительных устройств, обеспечивающих обработку сигналов, могут применяться различные цифровые устройства – от комбинационных схем из логических элементов до разнообразных процессоров, микроконтроллеров, программируемых логических матриц (ПЛИМ). Выбор того или иного вычислительного устройства обуславливается спецификой конкретной задачи, в частности требованиями к габаритным размерам, энергопотреблению и вычислительной мощности системы обработки.

В случае если разрабатывается система, которая будет производиться большими партиями, имеет смысл разработать для нее специализированный вычислительный блок, наиболее полно соответствующий запросам потребителя. В случае же мелкосерийного и опытного производства такой подход не окупается, и более удобно применить уже существующие элементы – универсальные процессоры, цифровые сигнальные процессоры либо ПЛИМ. Наибольшую гибкость в разработке обеспечивают универсальные процессоры, благодаря наличию расширенной системы команд, доступных сред разработки и инструментальных средств отладки, а также благодаря доступным программным решениям, которые были разработаны ранее для сходных проектов. Вместе с тем, применение специализированных процессоров и ПЛИМ может обеспечить более высокое быстродействие и энергоэффективность.

В современной вычислительной технике повышение быстродействия и энергоэффективности обеспечивается чаще всего двумя путями – использованием специализированных вычислительных блоков и параллельной обработкой данных. Специализированные вычислительные блоки служат для выполнения часто встречающихся ресурсоемких операций, например, вычисления тригонометрических функций или преобразования Фурье. Применение вычислительных систем с параллельной обработкой данных ставит дополнительную задачу при разработке алгоритмического и программного обеспечения обработки сигналов в радиотехнических системах. Для эффективного использования имеющихся вычислительных ресурсов необходимо разделить процедуру обработки на ряд независимых операций, которые могут проходить одновременно. При этом параллельная обработка позволяет существенно повысить скорость многих процедур обработки сигналов, что обуславливает растущий интерес к этой отрасли и находит свое отражение в публикациях [1 – 5]. В таких системах параллельной обработки сигналов могут применяться многоядерные универсальные процессоры, ПЛИМ, а также графические процессоры [4 – 6]. Применение персонального компьютера (ПК) либо кластера в роли вычислительного устройства РТС позволяет реализовать гибкую, высокопроизводительную и относительно недорогую систему цифровой обработки за счет применения серийно выпускаемых комплектующих и имеющихся средств разработки программного обеспечения.

Постановка задачи

В работах [4, 5] рассматривался вопрос об оценке вычислительных затрат процедур обнаружения и оценки параметров сигналов. При этом оценивалось среднее количество времени, затрачиваемое на выполнение полного цикла обработки сигнала на данной вычис-

лительной системе. При постановке задач имитационного моделирования и апробации методов и алгоритмов обработки сигналов такой подход является вполне приемлемым, так как дает возможность оценить общие затраты времени на выполнение запланированного объема задач. В то же время, при использовании ПК в роли вычислительного блока реальной РТС представляют интерес также вопросы, связанные с возможностью реализации обработки сигналов в реальном времени. В частности, необходимо оценить закон и параметры распределения времени выполнения задачи.

Подобный вопрос возникает из-за того, что ПК чаще всего работает под управлением операционной системы (ОС), не гарантирующей время отклика конкретной прикладной программы на воздействие. Помимо пользовательских программ (одна из которых в нашем случае реализует обработку сигналов) ПК также исполняет ряд системных процессов, обеспечивающих взаимодействие программной среды с аппаратурой, а также сервисные функции операционной системы. Диспетчер процессов ОС выделяет каждому процессу определенные временные интервалы; их величина и порядок выполнения процессов, как правило, не поддаются жесткому контролю со стороны пользователя ПК.

Основные положения работы

Для анализа производительности выбраны две процедуры, используемые в радиотехнических системах извлечения информации – обнаружение сигналов и измерение их параметров [4, 5], в дальнейшем обозначаемые как «задача 1» и «задача 2». Конфигурации тестовых систем приведены в табл. 1.

Таблица 1

Номер	Процессор	Число ядер	Тактовая частота, МГц	Пиковая вычисл. мощность, ГФлоп/с	Относит. вычисл. мощность	Объем ОЗУ, Гб	Пиковая проп. спос. ОЗУ, Гб/с	Тип
1	Intel Q9400	4	2660	42.56	1	8	17.0	ЦП
2	GF119	48	1620	155.5	3.65	1	9.6	ГП
3	GF116	192	1200	461	10.8	1	20.8	ГП
4	GF110	480	1464	1405	33.0	1.25	152	ГП

Система 1 использует многоядерный универсальный процессор (центральный процессор, ЦП), системы 2-4 – графические процессоры (ГП) фирмы NVIDIA. Приведенная в таблице пиковая вычислительная мощность оценена для расчетов с одиночной точностью (IEEE754 single precision).

На рис. 1 и 2 приведен пример реализации времени выполнения и закон распределения времени выполнения для задачи 1.

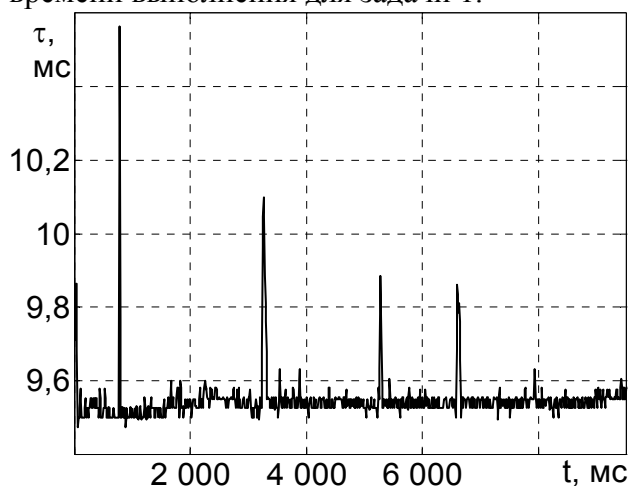


Рис. 1. Пример реализации $\tau(t)$

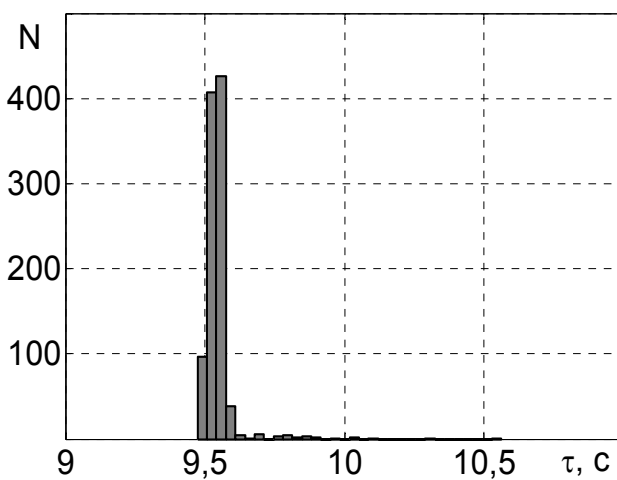


Рис. 2. Гистограмма зависимости $\tau(t)$

Анализ рис. 1 и 2 показывает, что наряду со случаями примерно постоянного времени выполнения (порядка 9.5 мс для приведенных данных) существует небольшой процент случаев, когда время выполнения процедур обработки оказывается существенно больше. Иными словами, гарантировать фиксированное значение времени обработки для системы на базе ПК оказывается невозможно. Однако применение ПК в качестве ядра системы реального времени возможно в тех случаях, когда можно задаться некоторой фиксированной вероятностью нахождения задержки в пределах заданного интервала.

Введем следующие обозначения. Пусть q_p - квантиль распределения времени выполнения вычислительной задачи для доверительной вероятности p . Тогда через

$$K_p = q_p / \bar{\tau} \quad (1)$$

обозначим коэффициент запаса – поправочный коэффициент, который необходимо учитывать для обеспечения гарантированного выполнения вычислительной задачи за заданное время. Иными словами, чтобы задача была выполнена за время $T \leq \bar{\tau}$ с вероятностью не менее p , необходимо увеличить вычислительную мощность системы обработки в K_p раз.

Таблица 2

Число потоков ЦП	Скорость выполнения
1	1
2	1.86
3	2.59
4	2.87

Таблица 3

Доверительная вероятность, %	ЦП, число потоков				ГП
	1	2	3	4	
90	1.10	1.11	1.11	1.16	1.01
95	1.11	1.13	1.25	1.19	1.02
99	1.11	1.14	1.28	1.57	1.03
99.9	1.18	1.39	1.52	6.15	1.05

С увеличением числа потоков скорость выполнения операций обработки возрастает (табл. 2, данные получены на системе 1). В то же время рост является нелинейным. Это обусловлено наличием в системе общих ресурсов, совместно используемых ядрами центрального процессора – прежде всего, оперативной памяти, шин и устройств ввода-вывода. В табл. 3 (система 1) приведены оценки коэффициента запаса. Этот коэффициент возрастает с ростом числа потоков ЦП, что может быть объяснено следующим образом. Действия, выполняемые операционной системой, также требуют участия ЦП. В то же время, если не все ядра ЦП заняты вычислительной задачей – то эти служебные действия могут быть выполнены свободными ядрами без ущерба производительности. Если же заняты все ядра – то выполнение служебных операций происходит за счет задач пользователя. Следует отметить, что для ГП характерны небольшие значения коэффициента запаса, что обусловлено отсутствием других системных или прикладных задач, использующих ресурсы ГП во время выполнения процедур обработки сигнала, а также спецификой данной вычислительной задачи. Дело в том, что для многоканального коррелятора характерно довольно большое число операций доступа к памяти, приходящееся на одну арифметическую инструкцию.

Объединим табл. 2 и 3, оценив оптимальную конфигурацию для системы на базе ЦП, реализующей обработку в реальном времени. Результаты приведены для системы 1.

При большой доверительной вероятности оптимальным вариантом для данной системы является использование трех вычислительных потоков при использовании 4-ядерного ЦП.

Таким образом, можно сформулировать следующий вывод: для систем, не рассчитанных на обеспечение режима жесткого реального времени, минимальные пиковые задержки выполнения обеспечиваются при условии, что имеется доступное ядро ЦП, берущее на себя служебные операции ОС.

Таблица 4

Доверительная вероятность, %	ЦП, число потоков			
	1	2	3	4
90	1.10	0.59	0.43	0.40
95	1.11	0.61	0.48	0.41
99	1.11	0.61	0.49	0.55
99.9	1.18	0.75	0.59	2.14

В табл. 5 и 6 приведены результаты аналогичных оценок для задачи 2 (системы 1,2) [4].

Таблица 5

Число потоков ЦП	Скорость выполнения
1	1
2	1.99
3	2.85
4	3.47

Таблица 6

Доверительная вероятность, %	ЦП, число потоков				ГП
	1	2	3	4	
90	1.00	1.00	1.00	1.12	1.02
95	1.01	1.01	1.02	1.20	1.03
99	1.02	1.01	1.07	1.68	1.12
99.9	1.03	1.02	1.26	2.65	1.38

Процедура параметризации спектров рассеянных сигналов более эффективно распределяется на большое число вычислительных ядер. Это связано с меньшим объемом памяти, совместно используемой разными потоками. Вместе с тем, коэффициент запаса для ГП в этой задаче возрос. По сравнению с задачей 1 для задачи 2 характерно большее число арифметических операций, приходящееся на один доступ к памяти, что обуславливает большую вычислительную нагрузку на арифметико-логические устройства ГП.

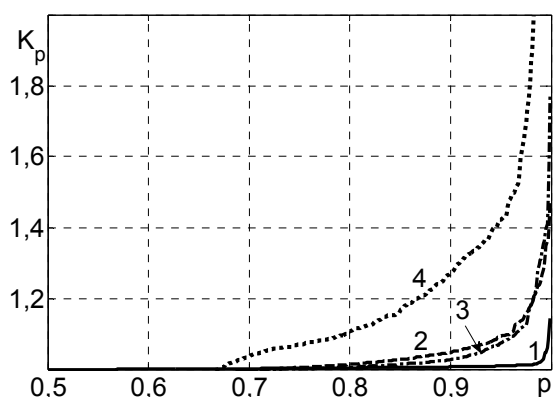


Рис. 3. Коэффициент запаса для задачи 2 (ЦП, 1-4 потока)

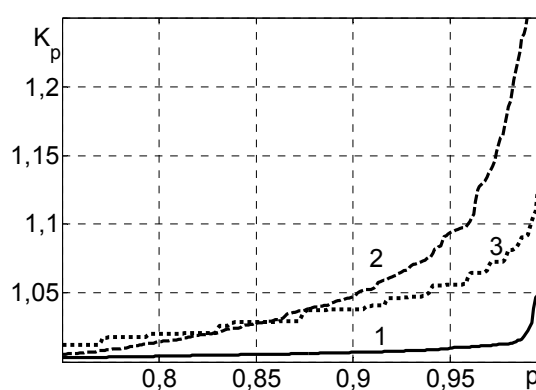


Рис. 4. Коэффициент запаса для задачи 2, ЦП, 1-2 потока (1,2), ГП (3)

С ростом доверительной вероятности, коэффициент запаса для всех конфигураций увеличивается. При этом наиболее медленными темпами он возрастает для ЦП при малом числе исполнительных потоков (рис.3, кривые 1-3), а также для ГП (рис. 4, кривая 3). Для случая, когда все ядра ЦП уже заняты (рис. 3, кривая 4) коэффициент запаса возрастает наиболее быстро.

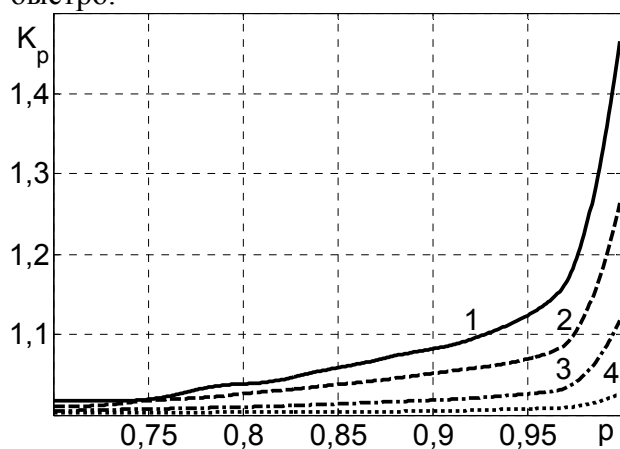


Рис.5. Коэффициент запаса для задачи 2 (1 – 64 точки, 2 – 128, 3 – 256, 4 – 512), доверительная вероятность 95%

На рис.5 приведены оценки коэффициента запаса для системы 4 для различных размеров выборки. Число арифметических операций возрастает пропорционально квадрату размера выборки, а объем пересылаемых данных – пропорционально размеру выборки. Таким образом, с ростом размера выборки увеличивается отношение времени арифметических расчетов к времени операций ввода-вывода. При этом коэффициент запаса уменьшается, что можно пояснить большей предсказуемостью времени расчетов за счет полной загрузки всех исполняемых блоков АЛУ графического процессора.

Таблица 7

Доверительная вероятность, %	ГП		
	с.2	с.3	с.4
90	1.733	1.023	1.038
95	1.735	1.035	1.056
99	1.738	1.161	1.093
99.9	1.739	1.501	1.218

В табл. 7 приведены оценки коэффициента запаса для всех конфигураций с графическими процессорами (системы 2,3,4) для размера выборки 512 точек. Для системы 2 характерны наибольшие значения коэффициента запаса, что можно объяснить тем, что данный ГП оснащен наиболее медленной памятью из всех протестированных систем.

Выводы

В вычислительных системах, не гарантирующих порядок выполнения процессов и распределение процессорного времени между ними, можно реализовать обработку сигналов, задавшись статистическими характеристиками временных интервалов выполнения операций. При этом гарантировать заданное время обработки сигналов с заданной доверительной вероятностью можно за счет применения вычислительного устройства с запасом по производительности, либо за счет использования дополнительных процессорных ядер, не занятых обработкой сигналов.

При прочих равных условиях, коэффициент запаса оказывается меньше для задач с меньшим объемом обращений к памяти, для систем со свободными ядрами ЦП, а также для систем на базе графических процессоров. Это позволяет сформулировать следующие требования к системе, реализующей обработку сигналов в режиме, приближенном к реальному времени:

1. В системах с несколькими процессорами либо с многоядерным процессором рационально оставлять один процессор (или ядро), не задействованным процедурами обработки сигналов;
2. Для обеспечения обработки в реальном времени критично обеспечение стабильно малой задержки обращения к оперативной памяти\4
3. Для задач, поддерживающих возможность параллельного выполнения, целесообразно использовать графические процессоры. При этом также важен вопрос о задержке доступа в локальную видеопамять графического процессора, что обуславливает целесообразность выбора систем с наиболее быстрой видеопамятью при прочих равных условиях.

Данная работа проводилась в рамках программы NVIDIA Academic Partnership, и автор хотел бы выразить признательность корпорации NVIDIA и лично Антону Джораеву за предоставленное оборудование (видеокарты GeForce GTS450), а также поблагодарить канд. техн. наук Зубкова О.В. за конструктивное обсуждение материала, представленного в статье.

Список литературы: 1. Xiaoshu Si, Hong Zheng High Performance Remote Sensing Image Processing Using CUDA // ISECS, pp.121-125, Third International Symposium on Electronic Commerce and Security, 2010. 2. Fung et al. Mediated Reality Using Computer Graphics Hardware for Computer Vision // Proceedings of the International Symposium on Wearable Computing 2002 (ISWC2002), Seattle, Washington, USA, Oct 7-10, 2002, pp. 83-89. 3. E. Alerstam, T. Svensson & S. Andersson-Engels. Parallel computing with graphics processing units for high speed Monte Carlo simulation of photon migration // J. Biomedical Optics 13, 060504 (2008). 4. А.И.Литвин-Попович. Обработка радиолокационных сигналов в параллельных вычислительных системах // Литвин-Попович А.И. // Радиотехника. – 2011. – Вып. 166. – С. 165-172. 5. А.И.Литвин-Попович. Обнаружение сигналов в условиях априорной параметрической неопределенности / Литвин-Попович А.И. // Радиотехника. Всеукр. межвед. науч.-техн. сб. – 2012. – Вып. 168. – С. 16-21. 6. Информационные материалы Nvidia. Режим доступа: <http://www.nvidia.com/object/personal-supercomputing.html>

Харьковский национальный
университет радиоэлектроники

Поступила в редколлегию 17.02.2014