TECHNOLOGY INSIGHT

COMPUTER DESIGN

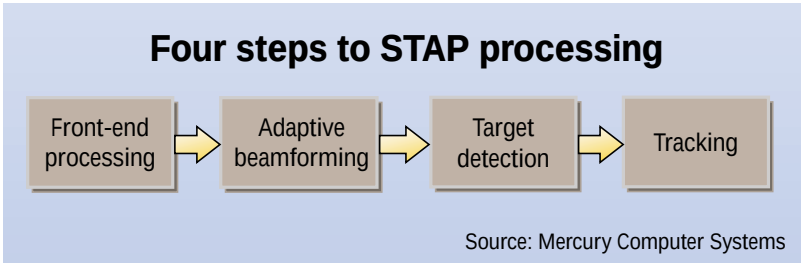
Blending processor types to overcome stream computing

– Craig Lund, principal technology consultant, Mercury Computer Systems

Combining a heterogeneous blend of processor types with a high-speed data pathway to allow information to travel between processing nodes is one way to solve the challenges of stream computing.

Space-time adaptive processing (STAP) radar is a good example of a digital signal processing (DSP) application that processes a vast data stream in real time. Mercury Computer Systems (Chelmsford, MA) recently delivered a heterogeneous system for processing STAP radar, based on the standard RACEway Interlink, to MIT's Lincoln Laboratory for use in the Navy's Mountaintop STAP program. This system comprises 948 Analog Devices (Norwood, MA) SHARC DSPs, and 24 Motorola (Austin, TX) PowerPC RISC processors, with peak performance of 118 Gflops and supporting memory of 20 Gbytes. These compute nodes reside in four separate chassis and are interconnected by the industry standard RACEway Interlink high-bandwidth data pathway.

For purposes of illustration (below), STAP processing can be divided into four logical steps, although in an actual STAP system, each of these steps consists of many processing operations. In the front end, arriving data undergoes a great deal of predefined, data-independent processing, such as pulse compression and Doppler filtering. This requires a large number of FFTs and complex vector dot products—functions well suited to DSP chips such as the SHARC processor, which can perform these operations efficiently, and which can be packed densely into the computer chassis.



The next step, requiring even more processing operations, is adaptive beamforming. This step consists of a number of very computationally intensive matrix operations. Again, because of the great volume of matrix operations in this step, DSPs' processing density makes them the component of choice.

RISC chips, such as the PowerPC, provide their best value in the tracking stage, acting on the data produced in the detection stage. In the tracking stage, processing follows code branches depending on the results of data analysis. (In comparison, earlier stages perform the same processing operations regardless of whether they detect open skies or an invading armada.) The amount of data entering this stage has been greatly reduced, so throughput is less of a consideration than earlier in the system. The RISC processor's superiority in running C code, combined with the lower processing density required in this phase, makes RISC the right chip for this final phase.

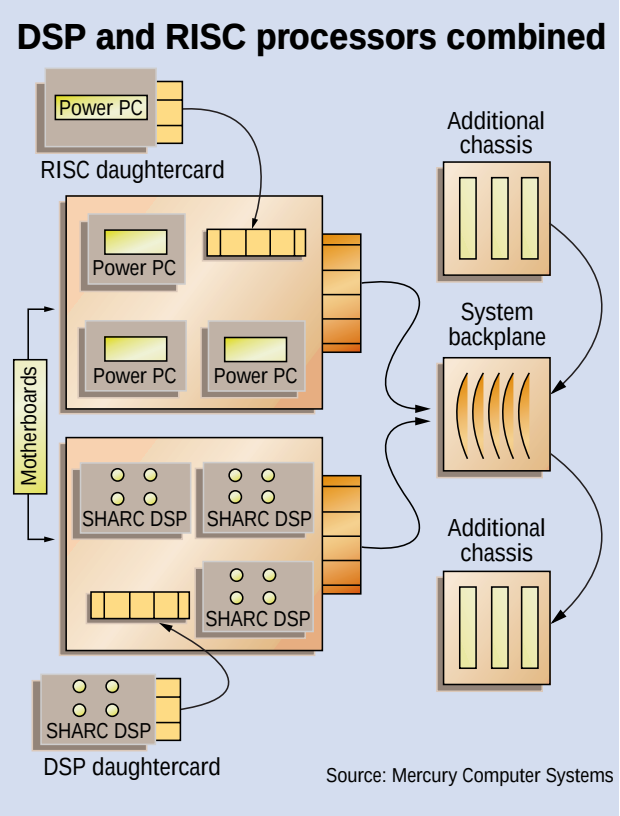
Even without DSP-like extensions, increases in RISC processor speed over the past 18 months raised this fundamental question: with RISC processors running so fast, why use DSP chips at all? Almost a year ago, some RISC chips had come to rival DSP chips in performing many signal-processing applications.

DSP chips still maintained a speed advantage, and provided a far greater processing density than RISC chips. (Processing density is defined here as MFLOPS of computational capacity per cubic foot and per watt of power, and is a key consideration in systems that will operate in a restricted space, such as aboard an aircraft.) In Mercury's RACE computer systems, three SHARC chips can fit in the real estate occupied by a single RISC processor, effectively providing a threefold increase in processor density.

Path toward convergence

This summer, Motorola announced its PowerPC 750 microprocessor running at speeds above 300 MHz. Although its efficiency (as measured by FLOP per cycle) doesn't match the SHARC chip, its faster clock speed and improved efficiency makes it suitable for typical vector operations. The table summarizes the pro-gression in RISC processor speeds over the past two years.

Reprinted from the November 1998 edition of **COMPUTER DESIGN**



Heterogeneity lets you combine DSP and RISC processors in a single logical system.

This comparison is based on measured performance on the complex vector dot product function using Mercury's Scientific Algorithm Library (SAL). The SHARC 21060 lags last year's PowerPC 603e, and this year's PowerPC 750, in terms of absolute performance. But in measures of floating point operations per cycle, its remarkable 1.86 FLOPS/cycle is far superior to either of these two RISC chips.

Recent announcements may result in the RISC chip overcoming even this difference. The new ADSP-21160 SHARC processor is expected to double the previous chip's flops-per-cycle rating achieving roughly 3.7 for complex vector dot product. At the same time, the G4 implementation of Motorola's AltiVec reportedly offers up to a fourfold increase over the PowerPC 750, or up to 4.88 FLOPs per second. If accurate, these numbers represent the first time a RISC processor has surpassed a DSP chip of the same generation in this metric.

Pressure to perform

The SHARC still maintains an advantage in performance density, but with its low-core voltage of 1.9 V, the new PowerPC 750 is ideal for the embedded systems market, where space and heat dissipation are at a premium. This introduction puts even greater pressure on the DSP community to respond with a chip that continues the specialty processor's dwindling advantages.

In addition, Motorola's new AltiVec technology represents another big stride in closing the performance gap between RISC microprocessors and DSP chips. This performance gap has stood as the main difference between the two architectures, since DSP chips historically boasted more computation units and achieved better utilization of data paths.

AltiVec overcomes the DSP architecture's lead in these key areas. Just as significantly, it does so with less impact on programmer productivity. The first AltiVec implementation can operate on up to 32 data sources with 16 destinations, in parallel, with a one-cycle throughput and often one-cycle latency.

Like a DSP chip, programming with AltiVec still involves C language extensions. Any non-standard extension comes with an impact on productivity today and portability tomorrow. AltiVec compiler extensions, however, are superior to those found in the DSP world because they represent additions to a scalar compiler, which are more advanced and mature than those found in the DSP world. AltiVec programmers should never need to drop into Assembly code to handle critical loops, avoiding a necessity that often bogs down DSP programmers.

AltiVec's data stream touch (DST) instructions bring the DSP technique of double-buffering to the PowerPC. DST puts four independent prefetch engines under program control. This allows the chip to perform one computational operation while simultaneously loading data for another.

In DSP programming, this operation is accomplished using internal RAM loaded under program control, using a DMA engine. But this technique doesn't address the need to keep coefficients inside the chip, since real-time data streams quickly displace coefficients from traditional caches. AltiVec's new LRU and transient options on load and store instructions satisfies this requirement by protecting data already in the L1 and L2 caches from replacement by newer data.

If history is an indicator of the future, chip manufacturing processes will likely continue to improve, and transistors will become even smaller and cheaper. DSP makers will continue using some of these transistors to make their products easier to program—a trend demonstrated by planned DSPs in Analog Devices' SHARC and Texas Instruments' (Houston, TX) TMS320Cxx product lines.

DSPs will also continue to find applications in high-volume devices, such as cellular telephones or disk drives. These applications present a number of mandatory calculations in situations where using the minimum amount of power or transistors is important, either to extend talk time, reduce manufacturing costs, or both. Because these are volume products, it becomes practical to invest in programming DSP applications to save five or 10 cents per unit manufactured.

DSP functionality will also be required in the next generations of

	SHARC 21060	PowerPC 603e	PowerPC 750
Clock rate	40 MHz	200 MHz	300 MHz
Performance	74 MFLOPs	137 MFLOPS	366 MFLOPS
FLOP/Cycle	1.86	0.68	1.22

Measured performance of dot product of two complex vectors of length 512 (function call overhead included).

signal and image processing computer systems. The only question is, in what form will those DSP functions be delivered? As process improvements drive RISC microprocessor prices lower, and DSP-like extensions trickle down from the high-end multiprocessor market, RISC microprocessors will offer increasing amounts of DSP functionality. Still, it's likely that DSPs optimized for particularly difficult functions will continue adding significant value to multicomputer systems.

Ultimately, the growth of field-programmable gate arrays (FPGAs) may have a major impact in this space. In contrast to DSP chips, which are optimized for particular signal processing operations, FPGAs will grow to provide arrays of transistors that can be configured according to the needs of system designers.

In some ways, this reflects today's trend in DSP manufacturing, in which DSP technology is most often sold as intellectual property (IP) for inclusion in ASIC cores. By their very name, ASICs are "application specific," but FPGAs will provide the ability to deploy DSP techniques that suit a specific purpose, then page in new configurations for entirely different applications.

The first implementation of AltiVec, in a chip Motorola hasn't officially named, but which is publicly referred to as G4, creates a new hybrid—the high-performance DSP. Real convergence will begin if future AltiVec implementations move down market, into integrated microcontrollers and eventually into ASIC cores. Motorola has prepared for this by creating a programming model that supports subsets targeted at specific DSP applications.

Motorola's programming model supports varying SIMD widths and software simulation of missing instructions (if, for example, floating point was dropped in a future derivative).

Upcoming generations of stream computers will likely be centered on one chip type, either a RISC microprocessor with DSP capabilities, or a DSP chip with RISC-like ease of programming and control function. For bottleneck functions, it's likely system designers will incorporate heterogeneous accelerators, such as FFT chips or FPGAs configured to perform a specific processing task.

Quite possibly, FPGAs configured around a DSP chip with RISC-like functionality, or a DSP-enabled RISC microprocessor, will allow system reconfiguration on an as-needed basis. Imagine a multicomputer system comprised of a thousand or more FPGAs with a sprinkling of multiprocessors, all interconnected by means of a high-speed backplane such as the RACEway Interlink. Configuring such a system, tuning it to address unanticipated bottlenecks, or even redeploying it in a new role, would be as simple as loading a new program is today.

In such a world, all systems would be heterogeneous—or none would be, depending on your perspective on the problem; but each system would contain precisely the hardware combinations needed for optimal performance.

Craig Lund is an industry consultant who devotes much of his time to Mercury Computer Systems. Mr. Lund is principal technology consultant, and is responsible for Mercury's longer range technology vision. He also acts as principal investigator on Mercury's research programs. Mr. Lund was founder of the IEEE POSIX family of standards. He holds a BSEE from the University of Connecticut.