



Symmetric

Multiprocessing

for Next-generation

Smartphone Platforms

By Steve Jahnke and Steve Krueger, Texas Instruments

Smartphone system designers can take advantage of a much wider range of power consumption/performance options with SoCs containing both symmetrical multiprocessing (SMP) for the high-level OS (HLOS) and specialized processors for media tasks. Legacy SMP architectures from the server/PC world are inadequate for mobile designs, however, and new SMP architectures are needed to accommodate the real-time and low-power requirements of mobile embedded systems such as smartphones.

In the past, fully implemented symmetric multiprocessing (SMP) has not played a significant role in smartphone design for several reasons, including:

- The absence of appropriate SMP software
- The difficulty of integrating multiple processors
- Smartphone power budgets
- Absence of low-power SMP processors

Today's smartphone applications also process data streams with algorithms that are primarily sequential. SMP is much more of a natural fit for systems with multiple data streams that must be processed in real-time as they arrive.

In addition to current smartphone capabilities such as SMS, music, video, email, web connected applications, and location-based services, next-generation smartphones will also require improved

performance, higher data throughputs over 3G and 4G networks, higher quality media, more sophisticated interfaces and applications that are yet to be imagined.

This combination will require more than just a leap in processing power. It will also require new multiprocessing architectures that extend battery life rather than diminish it and preserve the real-time performance of video, audio and media from a user's perspective.

Implications of Next-generation Mobile OSs

Web browsing with next-generation mobile OSs such as the Android platform will change smartphones substantially. In particular, web browsing will become more powerful by encompassing wide support of web data formats such as Flash; handling more sophisticated user interfaces; and, delivering much higher data rates for 3G and 4G. Multiple threading, such as that of Google's Chrome web browser -- already used in the PC world -- will be a big part of the enhancements. Multithreading is highly complementary to SMP operation and as these technologies make their way into the mobile world, multiple processes will run simultaneously on the smartphone SoC.

From a user's perspective, the ability to run multiple processes will result in performance, security and reliability improvements on an SMP system -- as long as a multi-threaded architecture is used. It is a challenge that mobile web browser developers are sure to take seriously. Smartphone chip designers and system designers can plan their next-generation products with a near certainty that SMP will become the operating mode for the mainstream smart phone architectures over the next few years.

In fact, the release of the ARM®Cortex™-A9 MPCore multi-core processor cleared the way for Texas Instruments (TI) to deliver a smartphone chip with SMP. The ARM Cortex-A9 MPCore architecture supports multiple processors with cache coherence, low-power and high-performance, making it the first SMP technology suitable for smartphone designs.

Running multiple processes simultaneously will, however, pose major challenges for system architects. SMP has achieved most of its present success in the server and PC environments. But certain characteristics that served SMP well there are not suitable for battery operated systems such as smartphones. In the PC/server environment, for example, the tradeoff between performance and power consumption is tilted heavily toward performance. Power management consists mostly of turning down the processor clock when possible. On the other hand, battery life is so critical for mobile systems that TI and its partners have developed power management into a full-fledged, complex discipline of its own.

In addition, server-oriented SMP architectures do not place the same high value on close-to-real-time performance as most mobile embedded systems do. Smartphones need to be able to quickly power down processors that are not needed. But Linux, to cite one example, has only one way to power off a processor -- the hot-plug API. That API is used to prepare the system for the manual replacement of processor boards. This is inadequate for battery-powered, embedded devices that must turn SoC processor cores on and off to conserve energy and extend battery life. At the same time, the real-time audio/video performance expectations of end users must be met.

To accomplish these two missions, the process that is running on the core that is to be shut down must be migrated to another core almost instantaneously. Instead of hot plugging, a more appropriate term might be high-speed, inter-core process swapping (HSICPS).

Odd as it may seem, the SMP architect must be sure that when a process migration is initiated, it actually reduces power consumption. There is, after all, an energy cost involved in de-powering a core only to power it up later.

There is also a real-time consideration when migrating processes. Moving the process from one core to another introduces some latency. SMP-based SoC architects must devise a means of checking whether the latency will cause a dropped frame in video or some other user-oriented quality problem.

These requirements are not supported by PC/server SMP. The PC/server environment is a much simpler world: Processors (or processor cores) are always on and power consumption is managed by turning down the frequency when it is OK. Also problematic is the fact that load-balancing algorithms migrates processes between cores consistently. This causes non-deterministic overhead in the application that could impact real-time performance requirements.

Heterogeneous Multiprocessing

The first step in realizing the goals of maximizing battery life and providing extremely high peak performance is to choose the appropriate multiprocessing (MP) architecture. MP architectures that utilize only multiple versions of the same general-purpose processor core provide some value. For example, when the system is processing multimedia data, which does not thread well, a core can be turned off without any impact on performance.

On the other hand, integrating heterogeneous cores into a SoC and fully implementing an MP architecture gives chip designers a very flexible platform for both power management and performance. In its OMAP products, Texas Instruments has a long and successful history of designing, implementing and integrating a variety of cores into SoCs.

These include: General-purpose processor (GPP) cores such as ARM processors; special-purpose programmable processor (SPP) cores such as graphics and video accelerators and DSPs; and, highly specialized processor (HSP) cores to execute specific algorithms like cryptography.

An MP architecture that organizes numbers of these cores on a SoC is known as heterogeneous MP. It offers mobile embedded systems the following advantages over other MP architectures.

- HSPs are 10-20 times more power efficient than GPPs for many algorithms
- HSPs can deliver required performance with much lower MHz than GPPs
- SPPs fill the gap in power and performance between HSPs and GPPs
- Multiple cores can be powered off when the application does not require them and because of their specialization, the opportunity to de-power cores is frequent.

TI's recently released OMAP 4 platform is an example of an advanced heterogeneous MP architecture combined with Cortex-A9 MPCore dual-processor SMP architecture. It is shown in Figure 1 with specific functional blocks labeled as GPP, HSP, or SPP cores.

The OMAP 4 platform follows a direct evolutionary path from the OMAP 3 product family. Its significant enhancement is not the architecture's heterogeneous nature, which has been evolving in a range of TI products for years, but the fact that it fully implements SMP and resolves PC/server SMP legacy issues. The OMAP 3 solutions used a heterogeneous SOC with a single core in the CPU domain. The base software for power and performance management was fine tuned in the OMAP 4 solution, increasing its functionality to accommodate SMP.

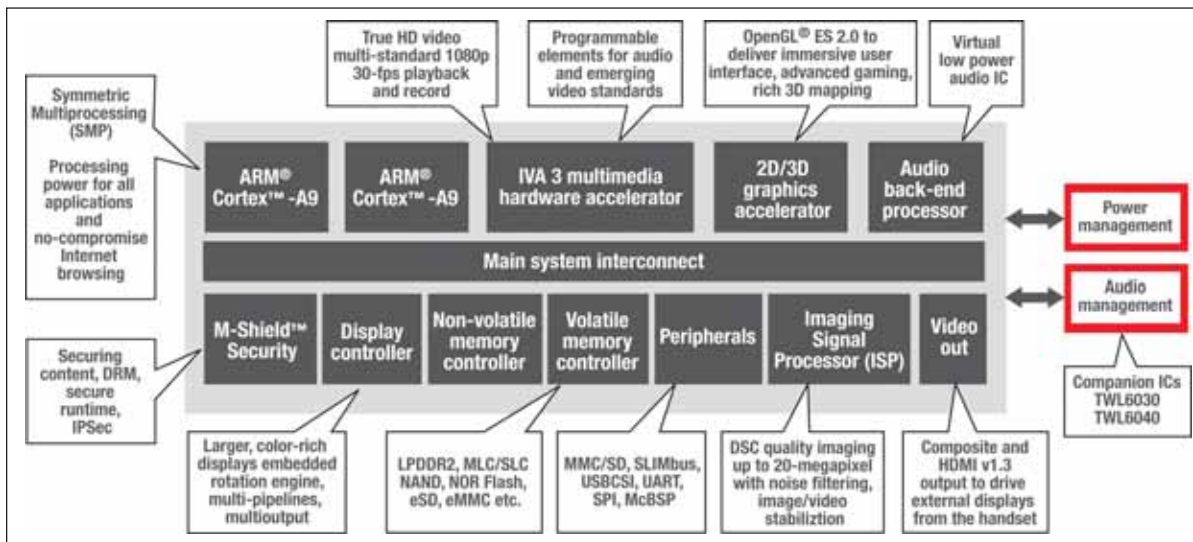


Figure 1: OMAP 4 processor leverages ARM Cortex-A9 MPCore

OMAP4's superior GPP performance comes from ARM Cortex-A9 MPCore processors running in a dual-processor SMP configuration. In a first for smartphones, the processors perform out-of-order processing for increased processing efficiency. Although high performance is available with both processors running, when demand declines one of the cores can be turned off to reduce power consumption.

The ARM wait-for-event (WFE) and wait-for-interrupt (WFI) instructions provide another means of reducing power when

fast response is required. Either of these instructions can put the CPU that executes it into a low-power wait state from which it can be awoken very quickly by an interrupt or event.

Finally, the ARM generic interrupt controller (GIC) handles the control and distribution of interrupts to the cores, providing a key part of the SMP solution.

Next-generation Smart Phones

The familiar video, imaging and audio applications of today's smartphones do not require multiprocessing themselves. The media that characterizes these applications, however, will flow in much greater quantities through the multiple browser tabs of next generation smart-phone OSs. As such, they will set the performance bar for each of those processes. Understanding the requirements of these use cases is critical to properly implementing SMP operations, including process swapping between cores.

In the OMAP 3 and OMAP 4 platforms, power management frameworks incorporate a policy manager and a resource manager to find the optimal operating configuration for the SoC. Use cases are created that allow the power management framework to mediate between the system architect's power management goals, the performance requirements of the use case, and quality metrics that are usually related to user

experience. To illustrate this concept, here are three common use cases and how they might constrain SMP operations such as shutting down a core, frequency scaling, or process swapping:

- Video will have a performance metric dependent on the selected resolution and the ability to display more than one video stream simultaneously. At the same time, the SMP architecture must meet a user-oriented quality metric, which in this case is the elimination of frame drops or stutters. Any action the SMP architecture might take in distributing available processing power to improve energy consumption, for example, cannot negatively impact the quality metric.
- Similarly, the smartphone's camera functionality use case sets its performance bar with the imaging chip's resolution and the computing power required by the post processing algorithms it implements.

The quality metrics are shot-to-shot delays and time required before the user takes the first picture.

- Audio performance is measured in playback time and its quality metric is eliminating audio drops.

A block diagram of a Policy Manager showing the main conceptual components is shown in Figure 2.

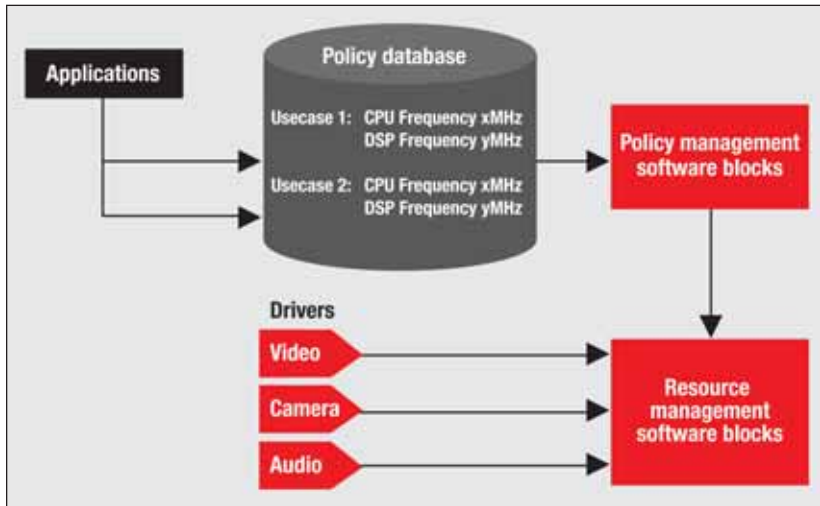


Figure 2: The Policy Manager determines the resource requirements for a use case and aligns them with existing hardware resources

A high-level request for resources from the application is referred into a policy database and an appropriate use case is selected. Based on the use case, the policy manager decides the SoC's appropriate power states (i.e. device frequency, which chip domains should be powered, etc.). The resource manager is responsible for determining if an actual power management resource can be placed in a certain state (i.e., if and how the clock should be scaled or turned off based on actual driver needs, etc.). In other words, the resource manager makes sure that the policy manager's requests can be implemented without a detrimental effect on lower level (such as driver) operations.

Use Cases and System Profiling Use cases stored in the policy database represent scenarios the system may encounter when processing various types of data.

An important new attribute in the OMAP 4 platform is a constraints database that helps the system decide if an un-utilized core should be turned off, or, if a process should be migrated from one core to another.

Constraints are created using a familiar technique called system profiling. Briefly stated, the system is bench tested using various scenarios for each use case. Test results determine whether constraints are generated for a particular use case. For example, the SMP load balancing algorithm may

determine that an under-utilized core should be turned off. But in checking the policy database, a constraint may be found stating that the core should remain active to service a real-time request that is likely to occur.

Constraints are implemented by adding an "is migratable" field to the policy database. Its value would allow process migration

to occur or not. Because some real-time applications such as a modem stack are extremely timing sensitive, the non-deterministic behavior of process migration between cores is not acceptable. For these use cases, a "no" would be entered in the field. In many other instances, the value would be "yes."

Another field has also been added to the policy database. It specifies the number of active cores for the use case. If the use case threads well or if high parallelism can be extracted, all cores would be active. If it is a single thread use case, this tag would be set to 1.

By adding these two fields, the OMAP 4 platform enables applications to force the number active cores desired and impose process migration constraints without specific directions from the application programmer. This is an important SMP attribute for widespread adoption.

The smartphone system architect, on the other hand, will have access to the Policy Database to run system profiling and create unique constraints. It remains to be seen how much value the new governor functionality will provide to any particular smart phone design. But having the functionality available as more and more high bandwidth content comes online can do no harm and might very well save the day for some designs.

Conclusion Thanks to the ARM Cortex-A9 MPCore, smartphones are beginning to deliver many advanced functions with enhanced user experiences, including web browsing, multimedia, instant messaging, WLAN connectivity and email. But smartphones are also endowed with mobile functions such as voice, SMS, Bluetooth, 3G and GPS. This combination of phone and computer functions is ushering in a new era of always-on, location-aware connectivity. SMP extends the smartphone experience into computer applications without sacrificing the traditional phone and media experiences. TI's OMAP 4 platform, based on dual-core ARM Cortex-A9 SMP provides industry-leading performance to deliver on this level of functionality.

END