MANIPULATING MEMORY IN SET-TOP BOXES: GETTING THE MOST FOR YOUR MONEY

Haig Krakirian
Pioneer Digital Technologies, Inc.

Abstract

Everyone knows that memory size in a set-top box does matter. What most people don't know is that memory configuration, type, and architecture are key factors in deploying successful digital services. Memory capability within a set-top is a complicated and varied proposition. Many customers of set-top boxes may not fully comprehend the critical deployment of multiple memory types and advanced application requirements when evaluating purchasing decisions. Quite simply, memory configuration in the box affects the performance and the price. We believe the decision-makers need to understand when the technology is worth their dollar.

We propose that readers may want to learn a bit more about memory models offered in set-tops before they conclude that the only difference in set-top boxes is the price. We intend to discuss the ways in which memory architecture can be designed for optimal efficiency and economy in a set-top box, why it matters to today's customer, and how it can affect tomorrow's service offerings.

DON'T FORGET ABOUT MEMORY

In the computing industry, memory capacity has always been driven by software demands. As software becomes more complex, the need for greater and greater amounts of memory escalates. This trend is similar but even more pronounced in the set-top space, however, because in addition to processing digital video, set-tops must also process ITV applications which are rich in multimedia content.

While there are many parallels between the technological advancements of the set-top and computer industries, the pace of development has lagged on the set-top side. Cost is a primary reason for this. MSOs are always on the lookout for that "sweet spot" between price and functionality. In order to keep set-top prices as low as possible, memory requirements traditionally have been based on the types of applications available at the time of development. For many MSOs, therefore, set-tops have been geared toward MPEG decoding and navigation software (including program guides), rather than for more advanced interactive applications. The only business model available to MSOs has been that of video delivery, making it hard to justify the extra cost for memory when there is no clear indication that there would be a return on that investment.

Meanwhile, software vendors who were hoping to find a dynamic new market in set-tops have been hesitant to commit resources to that market because they are not seeing adequate memory on which to base a decent application. This dearth of applications has further diminished the desire of MSOs to invest in memory because there simply is not enough to offer customers. In short, we're stuck in a classic chicken and egg scenario.

Fortunately, today's temporary hurdles are forcing the bar to be raised for both the application developer and the MSO. Rich content requires more color and animation, but it is also memory intensive. As in the computer industry, it will take cooperation and coordination between software developers and hardware manufacturers to bring about the compelling, eye-catching content that digital cable networks are capable of.

MEMORY OPTIONS

In order to understand the evolution of memory in set-tops, it is important to understand what types of memory are available and their functions. Today's memory types come in many forms and each has its own characteristics that can affect and improve the performance of set-top tasks. For set-top boxes, two primary types of memory are used: Random Access Memory (RAM), and Flash memory.

Random Access Memory or RAM, is memory that is available to the system processor when the set-top is booted up and running. RAM is essential to run and execute applications, but RAM memory loses its contents when the set-top is turned off, or if the content is not refreshed by an external charge. Thus, RAM is considered "readable/writable," meaning information, such as applications, can be written to, and read from its silicon, but not stored (without power), making it "volatile."

However, advances in RAM have blurred traditional definitions. The most common types of RAM today are Static RAM (SRAM) and Dynamic RAM (DRAM).

SRAM is faster than DRAM and accordingly more expensive. Because of the way DRAM is constructed, with individual cells composed of a transistor and capacitor, the contents must be continually refreshed by an external circuit to be retained in memory. SRAM requires no refreshing and retains data.

Because it's cheaper, DRAM is typically used in larger sizes. There are several different types of DRAM, including Synchronous DRAM (SDRAM), Double Data Rate RAM (DDR-RAM) and Extended Data Out RAM (EDO-RAM).

A large percentage of set-tops use SRAM as a digital channel receiving buffer, or the "interleaving buffer." SRAM serves this purpose well because the set-top processor needs fast access to RAM when processing video, but doesn't need to process a huge amount of data at once. On the other hand, DRAM is used for application execution, as a graphics rendering buffer and an MPEG decoding buffer.

RAM AND INTERACTIVITY

DRAM is quickly becoming the engine that caches and drives interactive applications in the set-top. Interactive TV applications are stored on servers and are broadcast, or streamed, over the network on specific channel frequencies. Set-tops "listen in" to those streams and begin caching, or saving, them in DRAM. Once the applications are fully loaded in DRAM, they are executed when a subscriber chooses to run that application.

One existing channel navigation paradigm is where each channel is an application. An email client is an example of an application that can be assigned to a channel allowing a user to simply go to a channel to check their email. In this case, when a user surfs over that channel, the application is loaded and executed out of DRAM.

Besides loading and executing applications, DRAM is used for loading data files that are used by those applications. A primary example is program guide data, which, because it changes daily, is best stored in DRAM. The set-top reads the program guide data files in much the same way it watches for interactive applications, and stores them in DRAM.

BACK IN A FLASH

The basic characteristic of flash memory is its "non-volatile" nature, meaning it can retain information without a backup battery or electric charge. When a box is turned off, flash memory will keep its data and content available for use when the box is turned back on. At the same time, flash memory can be re-programmed and erased, although a high voltage is needed to erase or make changes to the data stored in flash. Flash memory, then, is "persistent" in nature.

There are two types of Flash ROM, NAND type and NOR type. NAND is typically larger in memory size, but has slower access speeds.

Random access speeds of NOR can be from 70 ns to 100 ns per byte, and NAND can take 25 micro seconds to read 512 bytes. In contrast, page mode access, which allows quick read access from a specific range of memory, is generally faster in NAND. NOR can access at 20 ns to 25 ns per byte, but NAND can access at 50 ns per 512 byte. So, NOR can be faster than NAND for byte access, but NAND can be faster for bulk data access. Due to the nature of their individual capabilities, they are utilized for different tasks. NAND type is used for data storage, while NOR type is used for quick random access functions, such as program storage and execution.

Both NOR and NAND use FG (Floating Gate) cell technology. FG structure memory, especially in NOR type, may limit the potential growth of memory size because the memory cell is so complex. This is why Flash ROM is such an expensive element of a memory configuration. Fortunately a new type of Flash ROM called MONOS (metal oxide nitride silicon) type was recently introduced which allows for a more simplified cell structure, and therefore more potential for increases in memory size. Many manufacturers are betting on this technology to provide a high-memory, low-cost Flash ROM solution.

In the absence of hard drives in today's set-tops, flash memory has become the repository of the core applications of a cable video service, which include firmware, middleware, and a so-called "resident" application (comprised of a program guide, parental control, pay-per-view services, the core channel surfing application, etc.). In the event headend application servers go offline for any reason, subscribers will still be able to watch TV and change channels.

Over time, operators may determine that Web browsing is a critical application to their service offering, and placing the actual Web browsing application in flash memory will become more of an attractive option. Turning to this option, of course, means that more flash memory will be needed, beyond today's 4 megabyte (MB) standard.

MEMORY ARCHITECTURES

In addition to the types of actual memory available to set-top makers, there is a choice between memory architectures. A unified memory architecture uses the same memory to process video, graphics and other digital information. Unified has the advantage of providing a larger pool of available memory, depending on how many simultaneous functions the box is performing. If there is no MPEG video being processed, for example, there is more available memory for decoding graphics.

For applications developers, unified memory can be a drawback in that there is no easy way to know exactly how much memory will be available on a given set-top at a given time. This makes it tough to add the kinds of rich features desired by consumers because they push the envelope in memory usage. Unified memory can also be a drawback for the user who may start to see sluggish performance or scrolling hiccups if digital MPEG video is tuned, which takes priority in the decoding hierarchy.

A dual memory architecture solves the latter problem by providing graphics and video with their own dedicated memory. While this makes it impossible to take advantage of unused memory in, say, the video side of the box, it does allow developers a fixed memory limit when deciding what features to add to a particular application. The disadvantage to the set-top maker is cost. Duel memory cores require separate busses.

THE MARKET TODAY

The majority of advanced digital set-top boxes deployed today have a memory configuration of 4 MB of Flash and 8 MB of DRAM. For the MSOs who are deploying these boxes, the challenge is to utilize them to their fullest capacity.

With a 4/8 configuration, the operating system and core resident applications – such as the channel guide, pay-per-view, and video-on-demand – eat up about 5 MB of DRAM. That only leaves 3 MB of DRAM for the more advanced interactive applications, which is not a lot of memory when you're talking about advanced interactivity.

Compression techniques and intelligent caching techniques allow applications to take full advantage of a 3 MB footprint. However, decompressing the application data files can put a strain on CPU processing and bog the system down making a set-top unresponsive to user interaction. In addition, when you consider that any form of rich graphic content will require multiple 640x480 screens of at least 600 KB each, it becomes clear that 3 MB of memory can be eaten up in no time.

To be sure, there are a number of compelling applications that can run on 3 MB of memory. The question is which application should an MSO choose? MSOs are understandably hesitant to dedicate all of their remaining memory resources available in a set-top to just one application when their goal has always been to offer a host of interactive services as opposed to just one.

VOD IS KEY

Relying on the built-in MPEG video decoding capabilities of today's digital set-tops is most likely the solution to the limited memory problem. Interactive applications can leverage the capabilities of the current base of deployed digital set-tops to deliver rich and compelling content. After all, video is still king in the cable universe, and any new application will have to compete with it.

As long as cable operators are continuing their plant upgrades to provide streaming capability for every set-top, VOD will be the core technology for ITV applications. The cable infrastructure can be utilized to present rich, colorful video and animated content, rather than the more static web content. Best of all, VOD-streaming solutions do not require extra memory because you are using the video-decoding side of the box.

E-commerce, t-commerce, informational applications and other offerings would become more video-centric, rather than just text and pictures. A user could click through a few introductory pages and then get hit with a video stream. The PC has always had trouble with full-motion video, even under MPEG. The set-top has full-motion video built right in, so why not take advantage of that?

Most of the applications to-date have been data-centric, such as stock ticker information, sports scores and utility applications to manage programming options. There will have to be a re-thinking of the type of content presented and innovative re-use of graphics so as to offer a richer interactive experience from a minimal memory footprint. There is also a need to leverage the network more efficiently and manage content more wisely to cut down on the number of times the set-top has to access the information.

JUST AROUND THE CORNER

Clearly, then, the more advanced interactive content that requires larger chunks of RAM will have to wait for the next-generation boxes. But how long will it take? That is the million dollar question, but there are a number of factors that could propel the market a lot faster than most people think. High-definition television (HDTV) and personal video recorders (PVRs) will push the set-top to evolve much quicker than if current applications alone were leading the charge.

With PVRs, the caching of content requires a hard disk inside the box, which is memory that can be taken advantage of. The PVR may be the wedge that entices consumers into thinking of their set-tops as more than just dumb video terminals. After all, without a disk drive, web browsing on PC's would not be nearly as compelling as it is today.

HDTV could be another factor that propels additional memory in set-tops, mainly because it requires 32 MB of memory to decode. HD can be delivered via unified or dual memory architectures. Since most set-tops utilize unified memory, much of that capacity will not be used unless you are decoding HD.

Certainly, increasing set-top memory size is crucial to deploying advanced applications. While it is understood that memory is a scarce resource in set-tops today, there are a number of exciting developments right around the corner that may indeed accelerate the deployment of set-tops with expanded memory.

The key to success in the near term for MSOs is to understand current memory limitations and ways to leverage VOD technologies to offer enticing interactive services. It is also just as important to build a coherent roadmap for future interactive services based on the capabilities of next-generation set-tops just around the corner.

Haig Krakirian,
VP Software Engineering
Pioneer Digital Technologies, Inc.
2210 West Olive Avenue, 2nd Floor
Burbank, CA 91506
T 818.295.6628
F 818.295.6797
haig@pioneerdigital.com