Embedded Linux is upsetting the established order in the development tools marketplace. Developers are choosing Linux over traditional proprietary RTOS products for its superior reliability, for its cost effectiveness, and for its open source and standard APIs. After almost 20 years of using products like VRTX, pSOS, and VxWorks, developers are eager to embrace Linux, but express reservations about its ability to address real-time performance requirements. MontaVista Software has taken the lead within the embedded market by offering Linux as a Real-Time solution completely addressing concerns about Linux performance and showing how Linux can be applied to address a variety of real-time challanges.

REAL-TIME CHALLENGES IN LINUX

Off-the-shelf Linux presents three critical challenges for use in applications with real-time performance requirements.

1. LENGTHY BLOCKING TIMES

The Linux kernel regularly disables interrupts to protect critical sections of code. These sometimes lengthy interrupt-off times delay system response to external events, making system performance less deterministic. To aid in your design process, MontaVista characterizes Linux interrupt-off times, plus provides a comprehensive performance report for several configurations, as well as source code for the measurement tools used on the MontaVista Web site at www.mvista.com/realtime.

2. NON-PREEMPTIBLE LINUX KERNEL

Linux, in its present incarnation (2.2/2.4 kernel derivatives), is not a real-time operating system by traditional definition or functionality. To date, Linux design philosophy has emphasized throughput over responsiveness and determinism: its scheduler employs exhaustive fairness-based thread-selection, and can promote low priority threads to the head of the ready list, per its own criteria, to avoid starvation.
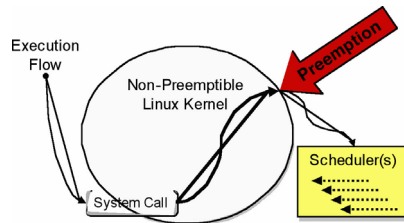

Figure 1. Non-preemptible Linux kernel

The Linux kernel, designed for optimized throughput, cannot be pre-empted. After a process execution enters the kernel via a system call, context switches will be deferred until the executing process blocks within the system call or exits the kernel. Because paths through the Linux kernel can vary greatly in length, this behavior dramatically limits user process responsiveness to external events.

3. EXHAUSTIVE, FAIRNESS-BASED SCHEDULING

The standard Linux scheduler inhibits predictable response times. To manage runnable threads, the scheduler maintains a linearly linked list of all executable processes. When the scheduler is called, it scans the entire list before completing process selection and initiating process dispatch. The linear nature of this algorithm causes scheduling time to increase linearly with the number of eligible threads. For small sets of executables, such overhead is acceptable. But as the number of threads grows, this algorithm contributes unacceptable delays in scheduler and system responsiveness.

MONTAVISTA SOFTWARE REAL-TIME SOLUTIONS: MONTAVISTA LINUX 2.1

FULLY PREEMPTIBLE LINUX KERNEL

Until recently, achieving a fully preemptive system (without breaking or stretching Linux APIs) was not deemed possible by the Linux community. Interestingly, there has for some time existed a preemptible kernel technology that resides at the core of the Linux kernel itself – the standard SMP (Symmetric Multi Processing) version of Linux. This version is extremely popular and broadly deployed in enterprise applications, offers highly preemptible and deterministic response when built and configured for uni-processor systems.

On mid-range Pentium systems, these kernels exhibit task response / preemption latencies averaging under 2.5 microseconds!

MontaVista Software builds and ships its standard MontaVista Linux kernel with user selectable options to build such a preemptible kernel for a variety of embedded applications, by applying the company's open source "preemption patch" to its standard configurations for x86/IA-32, PowerPC, StrongARM/XScale, ARM, MIPS, and SH architectures.

NOTE:

The MontaVista Linux preemptible kernel is available and shipping for all supported architectures: PowerPC, x86, MIPS, StrongARM, XScale, SH and ARM.

THE PREEMPTION PATCH

The MontaVista preemptible kernel patch modifies the definition (implementation) of the SMP kernel's primary IPC, the spinlock, changing it from its SMP specific implementation to a preemption lock. In the preemption scenario, as in its SMP application, the locking function acts as a control on reentrancy to critical sections of kernel code.

Additionally, the preemptible kernel patch modifies the interrupt handling software to allow rescheduling on return from interrupt if a higher priority process has become executable, even if the interrupted process was running in kernel mode (provided the process is not in a critical and locked region). SMP spin unlocks are also redefined to return the system to a preemptible state, and check if an immediate context switch is needed.

Lastly, the kernel build definition for a uniprocessor

target system is modified to include the spinlocks construct (implemented as preemption locks). Through these basic changes, the Linux kernel becomes generally preemptible (with short non-preemptible regions corresponding to the spinlocked regions in an SMP kernel). Process level responsiveness is dramatically improved, both on average and worst cases (See Figures 2 and 3).
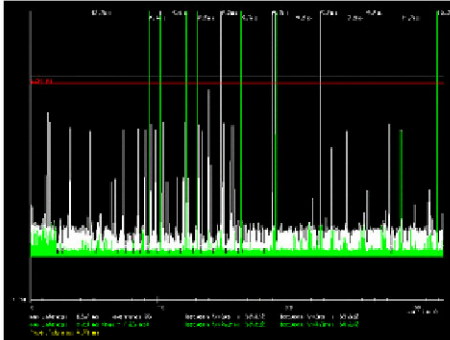

Figure 2. Audible glitching from a stock Linux 2.4 kernel under load while streaming audio.(In Figures 2 and 3 the red line represents audible effects)
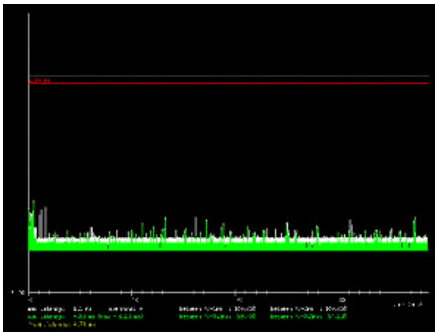

Figure 3. Minimal jitter from a preemptible Linux 2.4 kernel under load while streaming audio

THE MONTAVISTA REAL-TIME SCHEDULER (SEE FIGURE 4)
The MontaVista Real-Time Scheduler is a fully open, GPL-licensed source code module for Linux. This module for the Linux kernel addresses the serious process selection and dispatch time delays described in the previous section of this document.

FEATURES
• Enhances Linux system responsiveness
• Preserves standard Linux APIs and user thread-based programming model
• Eliminates the need for a second kernel "underneath" Linux
• Makes number of real-time priorities reconfigurable
• Addresses sub-millisecond scheduling requirements down to 150 microseconds

HOW IT WORKS
Executing ahead of the standard Linux scheduler, the MontaVista Real-Time Scheduler manages all threads marked as real-time (SCHED_FIFO and SCHED_RR policies). If there are no real-time policy threads, the MontaVista Real-Time Scheduler quickly and efficiently passes control to the standard Linux scheduler, with virtually no performance impact on standard Linux. It has a constant low overhead dispatch latency thereby eliminating the linear dispatch latency in the standard scheduler. The number of priorities is configurable between 99 and 2047 (default is 127).

MONTAVISTA REAL-TIME SOLUTIONS
MontaVista provides more than its Real-Time Scheduler. MontaVista's approach to providing real-time responsiveness with Linux also addresses other parts of the real-time Linux challenge.

OPTIMIZATION SERVICES
Since the company's founding, MontaVista has been applying its extensive embedded and real-time expertise to solve customer performance challenges including:
• Analyzing customer design requirements
• Debugging performance bottlenecks
• Benchmarking customer applications
• Tuning drivers and other system components

Contact Monta Vista for additional information on optimizing standard MontaVista Linux for real-time responsiveness and preemptible kernel technologies.

REAL-TIME CHARACTERIZATION TOOLS
Before MontaVista set out to provide an alternative Linux scheduler, our real-time engineering team performed extensive tests and experiments on standard Linux kernels for a variety of target CPUs and configurations.
MontaVista interrupt-blocking characterization data and tools are readily available from the MontaVista Web site at www.mvista.com/realtime.

Multiple options exist for making embedded Linux more deterministic in its response characteristics. With the advent of fully preemptible, open source kernel configurations/builds like those from MontaVista, developers with real-time needs and a desire to build their applications on an open platform like Linux, can enjoy enhanced determinism without giving up the open APIs and broadly supported core architecture that drew them to Linux in the first place.
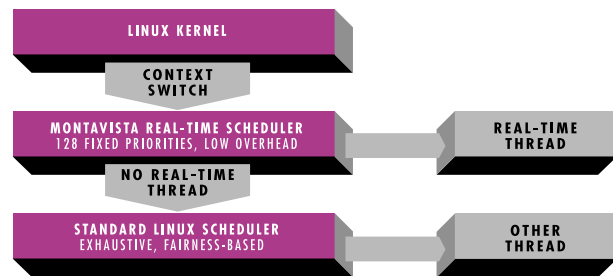

Figure 4. MontaVista Real-Time Scheduler

HIGH RESOLUTION POSIX TIMERS
MontaVista provides a solution for customers needing better timer resolution than that provided by the standard Linux 10ms tick. With our Carrier Grade Edition product, we provide timers with microsecond resolution and lower list overhead than standard timers.

MONTAVISTA™
SOFTWARE

tel: 408-328-9200
fax: 408-328-3875
www.mvista.com
info@mvista.com

Powering the Embedded Revolution