A Qualitative Comparison of MPSoC Mobile and Embedded Virtualization Techniques

Junaid Shuja, Abdullah Gani, Sajjad A. Madani

Faculty of Computer Science and Information Technology, University of Malaya, Malaysia
COMSATS Institute of Information Technology, Islamabad

Email: junaidshuja@siswa.um.edu.my, abdullah@um.edu.my, madani@comsats.edu.pk

Abstract—Virtualization is generally adopted in server and desktop environments to provide for fault tolerance, resource management, and energy efficiency. Virtualization enables parallel execution of multiple operating systems (OSs) while sharing the hardware resources. Virtualization was previously not deemed as feasible technology for mobile and embedded devices due to their limited processing and memory resource. However, the enterprises are advocating Bring Your Own Device (BYOD) applications that enable co-existence of heterogeneous OSs on a single mobile device. In this paper we investigate the processor architectures in the mobile and embedded space while examining their formal virtualizability. We also compare the virtualization solutions enabling coexistence of multiple OSs in Multicore Processor System-on-Chip (MPSoC) mobile and embedded systems. We advocate that virtualization is necessary to manage resource in MPSoC designs and to enable BYOD, security, and logical isolation use cases.

Index Terms—ARM, virtualization, hypervisors.

I. INTRODUCTION

Virtualization technology has been deployed in server space for decades since the early IBM mainframe systems [1]. In server space, virtualization enables hosting of multiple OSs on a single hardware platform. The hosted OSs are often homogeneous providing fault tolerance, resource management in terms of workload migration and consolidation, and energy efficiency. The Hypervisor or Virtual Machine Monitor (VMM) is placed between the hardware and host OSs and acts as a super OS. The hypervisor needs to efficiently share the hardware resources among guest OSs also known as Virtual Machines (VM). Moreover, hypervisor needs to translate guest OSs instructions if they do not meet the profile of underlying hardware. While translating guest OS instructions, virtualization adds the flexibility to execute non-compatible OS hardware combinations [2].

Mobile and embedded devices are resource constrained and OSs are reduced to minimal lines of code before deployment. Hosting multiple OSs on such resource constrained devices seems performance prohibitive. Mobile and embedded device such as Electronic Control Units (ECUs) can lose hard real-time functionality due to overheads of virtualization. However, due to recent advances in mobile technology such as multicore processor designs and higher memory volumes, virtualization has become feasible. There are three main use cases of virtualization in mobile and embedded systems: (a) virtualization enables BYOD scenario where an employee can execute both enterprise and personal applications on a single mobile device, (b) virtualization enables isolation of mixed criticality OSs on a device, and (c) virtualization leads to energy efficient resource utilization [3], [4], [5].

ARM based processors capture 90% of smartphone market [6]. However, there are a host of other processor architectures in the mobile and embedded space such as PowerPC [7], MIPS [8], and SH-4A [9] that are found in automobile controllers and consumer electronics. To implement virtualization efficiently, support from the underlying hardware platform is required such that: (a) the processor implements a hypervisor mode where the hypervisor can reside in a higher privilege level rather than occupying privilege level of the guest OS, and (b) each instruction that tries to change the context of hardware resources is privileged instruction [10]. Major mobile and embedded processor architectures including ARM lack the aforementioned hardware support for virtualization. Therefore, the hypervisor complexity increases as it needs to take care of privilege levels and instructions of the guest OSs.

Virtualization in mobile and embedded space is used to host heterogeneous OSs that address conflicting application requirements such as real time capability and user friendly GUI. Virtualization can enable concurrent hosting of a Real Time OS (RTOS) for sensor and communication control of the device and a general purpose OS (Android) for feature-rich GUI. IT enterprises usually control the technology their employees utilize during the office hours. This necessitates that each employee carry two mobile devices for enterprise and personal purposes. However, virtualization enables BYOD concept where both enterprise and personal OS and applications are hosted on a single mobile device in a logically isolated environment. Moreover, the BYOD concept significantly lowers the cost of hardware [11]. Embedded space devices often enforce strict security requirements on the hosted OS. However, virtualization enables concurrent execution of mixed criticality OSs such as a formally certified secure OS for automobile control and a general purpose OS for car entertainment system. Virtualization assures isolation of the domains by implementing separate address space for each hosted OS [12]. Furthermore, virtualization enables software reusability and survival of legacy software on newer platforms. It is estimated that most of smartphone devices are now shipped with a virtualization layer [13].

There a number of challenges pertaining to efficient virtu-
alization of mobile and embedded devices. Most of devices from mobile and embedded space have limited memory size. The hypervisor needs to maintain record of address mapping and instruction translation for the guest OSs in the memory, thus, over-burdening the resource. Moreover, MPSoC designs are often heterogeneous and aggregate processors with different instruction sets. Each processor architecture in the heterogeneous MPSoC design has different requirements for virtualizability. Moreover, most of OSs developed for mobile and embedded devices are designed to execute on single core. Hence, MPSoC designs require greater adoption and complexity in hypervisor design [14]. However, virtualization is necessary to manage resources in trending MPSoC architectures. Hardware support of virtualization in mobile and embedded devices is still in its early stages. Devices already deployed in smartphones and consumer electronics lack hardware support features for virtualization. Software enabled virtualization techniques utilized in such cases are not as efficient and consumes extra memory resources [15].

In this article, we provide a comparative analysis of software and hardware enabled virtualization techniques in mobile and embedded devices with particular focus on virtualization solutions addressing MPSoC designs. In section II, we provide a taxonomy of software and hardware based mobile and embedded virtualization techniques. In section III, we address the challenges and issues to virtualization of MPSoC mobile and embedded device. Section IV provides detailed comparison of various MPSoC virtualization techniques present in mobile and embedded devices and lists research challenges and issues to MPSoC devices. Section V concludes the discussion with future directions.

II. BACKGROUND

In the forthcoming subsections, we describe the formal and informal requirements of virtualization that are necessary for understanding of taxonomy of virtualization techniques. Additionally, we provide a short taxonomy of mobile and embedded virtualization solutions.

A. Formal and informal requirements of virtualization

The instructions of an Instruction Set Architecture (ISA) can be categorized into three classes: (a) instruction that properly execute in privileged mode and trap in unprivileged mode are privileged instructions, (b) instructions that try to modify the context of the system state are sensitive instructions, and (c) all other instructions are innocuous. Based on categorization of processor instructions, there are three formal requirements for classic virtualizability of an ISA. Firstly, the interface provided by the hypervisor to guest OSs should be identical to that of the underlying hardware. Secondly, most of the guest OS instructions should directly execute over the hardware with a small subset of instructions trapping to the hypervisor. Lastly, the hypervisor should remain in control of all virtualized physical resources [10].

Aside from the aforementioned formal requirements of virtualization, the resource constrained mobile and embedded devices enforce informal requirements for efficient virtualization. Firstly, while hosting multiple OSs, the device should not compromise real time capability. For example, the automobile controller requires hard real time response in vehicle tracking and management systems. Secondly, the virtualization layer should be scalable to the number of guest OSs and underlying processor cores. Otherwise, the virtualization will add no performance benefit in comparison to a single core non-virtualized environment. Thirdly, virtualization should be non-intrusive and require minimum guest OS modifications. Fourthly, the hypervisor should provide formal proof of security and guest OS domain isolation [16].

B. Virtualization techniques

Virtualization techniques are broadly categorized into Type-1 and Type-2 virtualization [17]. Type-1 virtualization enables the hypervisor to run in more privileged mode than the guest OSs. In this manner, the hypervisor fully controls access to the hardware resources made by guest OSs. Type-2 virtualization techniques execute the hypervisor as an application of the host OS. As a Type-2 hypervisor executes over an OS, it has less control of underlying hardware and can not provide resource sharing efficiently among hosted OSs. However, Type-2 hypervisors can be light-weight as they can host stripped down OSs environments [18]. Figure 2 illustrates the difference between Type-1 and Type-2 virtualization techniques.

Type-1 and Type-2 Virtualization techniques can be further classified into full virtualization, para virtualization, and Dy-
The main challenges to sharing and virtualization of memory space in mobile and embedded devices are: (a) maintaining Shadow Page Tables (SPT) for multiple guest OSs, (b) memory space protection and isolation, and (c) Translation Lookaside Buffer (TLB) maintenance on context switches. In a virtualized environment, the virtual address of the guest OS requires two-level translation. Firstly, it is translated from guest virtual address to guest physical address by the guest OS, and then, guest physical address in translated to machine physical address required by hypervisor. The hypervisor has to maintain address by the hypervisor. The hypervisor has to maintain the complexity of paravirtualization. In case an ISA is not fully virtualizable, paravirtualization remains the only feasible solution. Lightweight paravirtualization avoids rigorous pre-patching of guest OS and translates the sensitive non-privileged instructions at runtime. Microkernel based paravirtualization techniques are also popular in mobile and embedded devices. Microkernel based paravirtualization techniques emphasize the size reduction of the kernel to a minimal set while moving OS services from kernel space to user space. Figure 1 depicts full, para, and DBT based virtualization techniques.
separate SPTs for each guest. Whenever, there is a change in guest OS page tables, the hypervisor has to capture and mirror that change in its SPT so that it is able to perform the second stage of address translation [28]. Similarly, the hypervisor has to maintain some form of page table tagging mechanism so that it can isolate memory space of virtualized guest OSs. The hypervisor has to emulate domain tagging mechanism, if domain tagging support is not provided in hardware, such as old versions of ARM processors. Moreover, whenever a context switch takes place between the guest OSs or the hypervisor, TLB needs to be flushed and then re-populated for the currently executing host. As such context switches occur frequently in the virtualized environment, TLB misses increase resulting in high memory latency [29].

Mobile and embedded devices consist of a plethora of sensors and communication boards such as WiFi, Bluetooth, and cellular radio. Sharing these components among multiple guest OSs requires fine grained timescale multiplexing. Moreover, GUI based devices also require consideration for GUI sharing and switching among multiple guest OSs. While the currently executed guest OS get full access of the GUI, the remaining guests have to wait for their turn [18]. Furthermore, special vector processors such as ARM Vector Floating-Point (VFP) requires virtualization support for resource sharing [30].

IV. COMPARISON OF MOBILE AND EMBEDDED HYPERVERSORS

There are a number of commercial and research based mobile and embedded hypervisors. We will focus on detailed comparison of those hypervisors that support MPSoC designs.

1) Virtual Hellfire Hypervisor: Virtual Hellfire Hypervisor [14] is a MPSoC virtualization solution based on MIPS processors and Hellfire guest OSs. The Virtual Hellfire Hypervisor is microkernel based and provides a hypercall interface for the guest OSs to execute privileged instructions. The memory is partitioned into fixed chunks at kernel load time based on number of guests OSs. The intra-cluster communication between CPUs in the same cluster is done through hypercall interface while accessing a shared memory region among two clusters. Inter-cluster communication over the Network-on-Chip (NoC) architecture is done via wrapper functions.

2) ITRI: ITRI [31] is an ARMv7 based hybrid hypervisor that is utilizes both paravirtualization and DBT based full virtualization. ITRI supports MPSoC design and multiple guest OSs in form of paravirtual Linux 2.6 kernels. The ITRI hypervisor is based on a modified version of Linux based Kernel Virtual Machine (KVM). The guest Linux kernel is patched by removing unnecessary privileged instructions and redundant context switches. The paravirtualization of Linux kernel helps reduce trap-and-emulate and TLB maintenance cost. QEMU emulator is used to run guest OS instances along with real and backend drivers. ITRI is one of few mobile hypervisors that supports guest migration [34]. The performance evaluation of native and VM based execution show negligible overhead in compute intensive tasks while five times overhead in worst case for I/O intensive tasks.

3) KVM/ARM: KVM/ARM [26] is light weight paravirtualization solution based on MPSoC ARM design. The light weight paravirtualization uses scripts based on regular expressions to replace sensitive non-privileged instructions with trap-and-emulate procedure. Inter-processor Interrupts (IPI) are used as a communication mechanism between multiple processors. In this context, hardware extensions for virtualizing the generic interrupt controller (GIC) are utilized. A virtual GIC interface is introduced for each CPU with a corresponding hypervisor control interface. Interrupt can be raised by writing to the list registers in the hypervisor control interface. Resultantly, the hypervisor directly raises hardware interrupt to the corresponding guest OS kernel without causing a trap. The performance evaluation shows 10% overhead for realtime applications as compared to native execution.

4) Proteus Hypervisor: Proteus [25] is a PowerPC architecture based hypervisor that supports MPSoC designs. Proteus is a hybrid hypervisor that implements both full and paravirtualization without specific hardware support. Although hardware support for virtualization was released in newer PowerPC devices, Proteus is based on previous releases. Proteus implements a microkernel by moving I/O device drivers to user space while the supervisor space only executes hyper-call handler, VM scheduler, and inter-process communication modules. Proteus design is Symmetric Multiprocessing (SMP); when a guest traps to the hypervisor, the hypervisor takes control of its assigned core. A small part of the hypervisor runs on each core to support full virtualization and forward inter-processor interrupts. Proteus hypervisor implements a semaphore based solution for access to shared devices in the user space. Experimental results show that virtualization overheads for most of instructions are below hundreds of processor cycles.

5) AUTOSAR based Hypervisor: Reinhardt et al. [32] proposed a hypervisor design for security critical automotive devices based on AURIX memory management less (MMU) processor and AUTOSAR guest OS. Automotive devices require domain and fault isolation between the guest OSs. Moreover, the number of guest OS and there memory requirements are fixed in such environments, allowing for fixed partitioning of memory and doing away with MMU. Instead Memory Protection Unit (MPU) is utilized for guest OS domain isolation. The hypervisor sets up itself and Virtual Device Emulators (VDEs) in the highest privileged mode while allocating accessible memory regions for each guest OS. The proposed hypervisor uses paravirtualization to emulate privileged instructions. Interrupts are first handled by the hypervisor and then forwarded to the corresponding guest OS through its vector table. The Microcontroller Abstraction Layer (MCAL) of AUTOSAR is ported to accommodate hypervisor startup code, guest OS memory allocations, and VDE setup for inter-VM communications. The paravirtualized solution results in large overheads over native executions.

6) RAMPSoCVM: Runtime Adoptive Multiprocessor System-on-Chip Virtual Machine (RAMPSoCVM) [33] utilizes embedded Linux kernel to extend support for MPSoC virtualization with Message Passing Interface (MPI). As the hypervisor is based on embedded Linux kernel, it comes with
integrated support for multitasking and multithreading. As embedded Linux without BIOS detection, OpenFirmware is integrated as a module to detect dynamic MPSoC platform. The hypervisor manages the resources while meeting real-time constraints of the applications. To do so, CAP-OS is adopted as a module in the Linux kernel that services hard real-time constraints of applications. A MPI is provided to the applications to communicate with hypervisor services and device drivers managed by CAP-OS module. A Qualitative comparison of the discussed MPSoC mobile hypervisors is provided in Table [I].

All the above discussed hypervisors have two main characteristics in common: (a) they are specific to mobile and embedded device space and based on such processor architectures, and (b) they are based on MPSoC architectures. However, despite of these studies, open research issues remain for future considerations. Most of the researchers have ignored the issue of heterogeneous MPSoC architectures, where different processor architectures enforce different requirements for virtualization. Although paravirtualization can enable virtualization of heterogeneous OSs, hardware supported full virtualization is desirable due to its efficiency. To the best of our knowledge, their is no such solution available in research and commercial virtualization techniques. Memory optimization for DBT based full virtualization is also not explored in the aforementioned virtualization solutions. While user responsiveness is driving mobile devices to higher power and enabling concurrent execution of multiple OSs, embedded devices still lack sufficient processing power and memory to support virtualization. Moreover, embedded devices often have hard real-time constraints that might be compromised in a resource sharing environment. Security and isolation of multiple domains can be addressed by formal verification of the hypervisor. However, such formal verifications are scarce in literature [35].

V. CONCLUSION

In this paper, we carried out qualitative analysis of MPSoC design based mobile and embedded device virtualization solutions. We assert that the virtualization of mobile and embedded devices needs to be inspected from different aspects. As mobile devices such as smartphones and tablets are rapidly advancing in processor power and memory size, embedded devices are lacking behind. Most of the virtualization solutions are either paravirtual or DBT based full virtualization. Paravirtualization techniques require expensive guest OS patching while DBT based full virtualization solutions are inefficient in memory utilization. Hardware supported full virtualization is desired is such resource constrained platforms. Moreover, a performance analysis of paravirtualization, DBT based full virtualization, and hardware supported full virtualization techniques is yet to be debated. Furthermore, hardware supported full virtualization solutions for heterogeneous MPSoC designs need to be devised.

Acknowledgments

This work is partially funded by the Malaysian Ministry of Education under the High Impact Research Grant of University of Malaya UM.C/625/1/HIR/MOE/FCSIT/03.

REFERENCES

[1] G. J. Popek and R. P. Goldberg, “Formal requirements for virtualizable third generation architectures,” Communications of the ACM, vol. 17, no. 7, pp. 412–421, 1974.
[2] A. Aguiar and F. Hessel, “Embedded systems’ virtualization: The next challenge?” in 21st IEEE International Symposium on Rapid System Prototyping (RSP), 2010, pp. 1–7.
[3] J. Shuja, S. A. Madani, K. Bilal, K. Hayat, S. U. Khan, and S. Sarwar, “Energy-efficient data centers,” Computing, vol. 94, no. 12, pp. 973–994, 2012.
[4] G. Heiser, “The role of virtualization in embedded systems,” in Proceedings of the 1st workshop on Isolation and integration in embedded systems, 2008, pp. 11–16.
[5] J. Shuja, K. Bilal, S. Madani, M. Othman, R. Ranjan, P. Balaji, and S. Khan, “Survey of technologies and architectures for designing energy-efficient data centers,” Systems Journal, IEEE, vol. PP, no. 99, pp. 1–13, 2014.
[6] V. Do, “Security services on an optimized thin hypervisor for embedded systems,” Ph.D. dissertation, Faculty of Engineering LTH at Lund University, 2011.
[7] A. Mittal, D. Bansal, S. Bansal, and V. Sethi, “Efficient virtualization on embedded power architecture® platforms,” in ACM SIGPLAN Notices, vol. 48, no. 4, 2013, pp. 445–458.
[8] A. Aguiar and F. Hessel, “Virtual hellfire hypervisor: Extending hellfire framework for embedded virtualization support,” in 12th International Symposium on Quality Electronic Design (ISQED), 2011, pp. 1–8.
[9] W. Kanda, Y. Yamura, Y. Kinebuchi, K. Makijima, and T. Nakajima, “Spumone: Lightweight cpu virtualization layer for embedded systems,” in Embedded and Ubiquitous Computing, 2008. EUC’08. IEEE/IFIP International Conference on, vol. 1, 2008, pp. 144–151.
[10] N. Penneman, D. Kudinskas, A. Rawsthorne, B. De Sutter, and K. De Bosschere, “Formal virtualization requirements for the arm architecture,” Journal of Systems Architecture, vol. 59, no. 3, pp. 144–154, 2013.
[11] S. Earley, R. Harmon, M. R. Lee, and S. Mithas, “From byod to byoa, phishing, and botnets,” IT Professional, vol. 16, no. 5, pp. 16–18, 2014.
[12] P. Varanasi and G. Heiser, “Hardware-supported virtualization on arm,” in Proceedings of the Second Asia-Pacific Workshop on Systems, 2011, p. 11.
[13] Gartner Group. (2008, July) Forecast: Mobile devices, worldwide, 2003-2012., [Online]. Available: https://www.gartner.com/doc/731007/forecast-mobile-devices-worldwide.
A. Aguiar, F. G. de Magalhaes, and F. Hessel, “Embedded virtualization for the next generation of cluster-based mpsocs,” in 22nd IEEE International Symposium on Rapid System Prototyping (RSP), 2011, pp. 113–119.

D. Brash, “Recent additions to the armv7-a architecture,” in IEEE International Conference on Computer Design (ICCD), 2010.

G. Heiser, “Virtualizing embedded systems: why bother?” in Proceedings of the 48th Design Automation Conference, 2011, pp. 901–905.

J.-H. Ding, C.-J. Lin, P.-H. Chang, C.-H. Tsang, W.-C. Hsu, and Y.-C. Chung, “Armvisor: System virtualization for arm,” in Linux Symposium, 2012, p. 93.

K. Barr, P. Bungale, S. Deasy, V. Gyuris, P. Hung, C. Newell, H. Tuch, and B. Zoppis, “The vmware mobile virtualization platform: is that a hypervisor in your pocket?” ACM SIGOPS Operating Systems Review, vol. 44, no. 4, pp. 124–135, 2010.

Z. Gu and Q. Zhao, “A state-of-the-art survey on real-time issues in embedded systems virtualization,” Journal of Software Engineering and Applications, vol. 5, no. 4, pp. 277–290, 2012.

P. Varanasi, “Implementing hardware-supported virtualization in okl4 on arm,” Ph.D. dissertation, University of New South Wales, 2010.

C. Dall and J. Nieh, “Kvm for arm,” in Proceedings of the Ottawa Linux Symposium, Ottawa, Canada, 2010.

G. Heiser and B. Leslie, “The okl4 microvisor: Convergence point of microkernels and hypervisors,” in Proceedings of the first ACM asia-pacific workshop on Workshop on systems. ACM, 2010, pp. 19–24.

A. Suzuki and S. Oikawa, “Analysis of the arm architecture’s ability to support a virtual machine monitor through a simple implementation,” International Journal of Networking and Computing, vol. 3, no. 1, pp. pp–153, 2013.

ARM Architecture Group, “Virtualization extensions architecture specification,” ARM Limited, http://www.arm.com/products/processors/technologies/virtualization-extensions.php, Tech. Rep., 2010.

K. Gilles, S. Groesbrink, D. Baldin, and T. Kerstan, “Proteus hypervisor: Full virtualization and paravirtualization for multi-core embedded systems,” in Embedded Systems: Design, Analysis and Verification. Springer, 2013, pp. 293–305.

C. Dall and J. Nieh, “Kvm/arm: the design and implementation of the linux arm hypervisor,” in Proceedings of the 19th international conference on Architectural support for programming languages and operating systems, 2014, pp. 333–348.

C. Dall and J. Nieh, “Kvm/arm: Experiences building the linux arm hypervisor,” 2013.

R. Mijat and A. Nightingale, “Virtualization is coming to a platform near you,” ARM White Paper, 2011.

J.-Y. Hwang, S.-B. Suh, S.-K. Heo, C.-J. Park, J.-M. Ryu, S.-Y. Park, and C.-R. Kim, “Xen on arm: System virtualization using xen hypervisor for arm-based secure mobile phones,” in 5th IEEE Consumer Communications and Networking Conference, 2008, pp. 257–261.

S. Yoo, S.-b. Yoo, and C. Yoo, “Virtualizing arm vfp (vector floating-point) with xen-arm,” Journal of Systems Architecture, vol. 59, no. 10, pp. 1266–1276, 2013.

A. Smirnov, M. Zhdiko, Y. Pan, P.-J. Tsao, K.-C. Liu, and T.-C. Chiu, “Evaluation of a server-grade software-only arm hypervisor,” in Proceedings of the IEEE Sixth International Conference on Cloud Computing, 2013, pp. 855–862.

D. Reinhardt and G. Morgan, “An embedded hypervisor for safety-relevant automotive e/e-systems,” in Industrial Embedded Systems (SIES), 2014 9th IEEE International Symposium on. IEEE, 2014, pp. 189–198.

D. Gehringer, S. Werner, M. Hubner, and J. Becker, “Rampsocvm: runtime support and hardware virtualization for a runtime adaptive mpsoc,” in Field Programmable Logic and Applications (FPL), 2011 International Conference on, 2011, pp. 181–184.

S. Groesbrink, “Virtual machine migration as a fault tolerance technique for embedded real-time systems,” in Software Security and Reliability-Companion (SERE-C). 2014 IEEE Eighth International Conference on, 2014, pp. 7–12.

G. Heiser, J. Andronick, K. Elphinstone, G. Klein, I. Kuz, and L. Ryzhyk, “The road to trustworthy systems,” in Proceedings of the fifth ACM workshop on Scalable trusted computing, 2010, pp. 3–10.