World Academy of Science, Engineering and Technology
International Journal of Computer and Systems Engineering
Vol:11, No:03, 2017

A Bottleneck-Aware Power Management Scheme in Heterogeneous Processors for Web Apps

Authors : Inyoung Park, Youngjoo Woo, Euiseong Seo

Abstract : With the advent of WebGL, Web apps are now able to provide high quality graphics by utilizing the underlying graphic processing units (GPUs). Despite that the Web apps are becoming common and popular, the current power management schemes, which were devised for the conventional native applications, are suboptimal for Web apps because of the additional layer, the Web browser, between OS and application. The Web browser running on a CPU issues GL commands, which are for rendering images to be displayed by the Web app currently running, to the GPU and the GPU processes them. The size and number of issued GL commands determine the processing load of the GPU. While the GPU is processing the GL commands, CPU simultaneously executes the other compute intensive threads. The actual user experience will be determined by either CPU processing or GPU processing depending on which of the two is the more demanded resource. For example, when the GPU work queue is saturated by the outstanding commands, lowering the performance level of the CPU does not affect the user experience because it is already deteriorated by the retarded execution of GPU commands. Consequently, it would be desirable to lower CPU or GPU performance level to save energy when the other resource is saturated and becomes a bottleneck in the execution flow. Based on this observation, we propose a power management scheme that is specialized for the Web app runtime environment. This approach incurs two technical challenges; identification of the bottleneck resource and determination of the appropriate performance level for unsaturated resource. The proposed power management scheme uses the CPU utilization level of the Window Manager to tell which one is the bottleneck if exists. The Window Manager draws the final screen using the processed results delivered from the GPU. Thus, the Window Manager is on the critical path that determines the quality of user experience and purely executed by the CPU. The proposed scheme uses the weighted average of the Window Manager utilization to prevent excessive sensitivity and fluctuation. We classified Web apps into three categories using the analysis results that measure frame-per-second (FPS) changes under diverse CPU/GPU clock combinations. The results showed that the capability of the CPU decides user experience when the Window Manager utilization is above 90% and consequently, the proposed scheme decreases the performance level of CPU by one step. On the contrary, when its utilization is less than 60%, the bottleneck usually lies in the GPU and it is desirable to decrease the performance of GPU. Even the processing unit that is not on critical path, excessive performance drop can occur and that may adversely affect the user experience. Therefore, our scheme lowers the frequency gradually, until it finds an appropriate level by periodically checking the CPU utilization. The proposed scheme reduced the energy consumption by 10.34% on average in comparison to the conventional Linux kernel, and it worsened their FPS by 1.07% only on average.

Keywords : interactive applications, power management, QoS, Web apps, WebGL
Conference Title : ICCS 2017 : International Conference on Computer Science
Conference Location : Osaka, Japan
Conference Dates : March 30-31, 2017