

Optimizing kernel config for perf-defconfig/Perf builds

Recommended kernel config file

It is recommended to use <chipset>-perf_defconfig as the user build reference configuration. It is tuned for better system performance and shows up as better kernel memory footprint.

For LA based Environment-

The corresponding perf-defconfig reference is declared in AndroidBoard.mk file as shown below.

The chipset specific AndroidBoard.mk file will have to be taken as reference.

For example, the perf-defconfig reference for msm8953, is explained below:

```
#-----  
-----  
  
# Compile Linux Kernel  
#-----  
-----  
  
ifeq ($(KERNEL_DEFCONFIG),)  
    ifeq ($(TARGET_BUILD_VARIANT),user)  
        KERNEL_DEFCONFIG:= msmcortex-perf_defconfig  
    else  
        KERNEL_DEFCONFIG:= msmcortex_defconfig  
    endif  
endif
```

- For any chipset, the corresponding AndroidBoard.mk reference file may be found under the following paths for MSM8953, MSM8909W.

device/qcom/<(chipset_name)_32/64>/AndroidBoard.mk

We can prepare image for userdebug build with perf-def config by making below changes, ex for msm8909w:

Go to device/qcom/msm8909w/AndroidBoard.mk

force use of KERNEL_DEFCONFIG as perf defconfig.

```
ifeq ($(KERNEL_DEFCONFIG),)
```

```
ifeq ($(TARGET_BUILD_VARIANT),user)
```

```
KERNEL_DEFCONFIG := msmXXXX-perf_defconfig
```

```
else
```

```
KERNEL_DEFCONFIG := msmXXXX-perf_defconfig
```

```
endif
```

```
endif
```

For LE based environment: -

APQ8053 : -

The corresponding perf-defconfig reference is declared in **apq8053.conf** file as shown below. The corresponding **apq8053.conf** reference file may be found under the following paths :- **“/poky/meta-qt-bsp/conf/machine/apq8053.conf”**

```
#Perf Kernel config
```

```
KERNEL_PERF_IMAGETYPE = “Image-dtb”
```

```
KERNEL_PERF_DEFCONFIG = “apq8053_IoE-perf_defconfig”
```

APQ8009: -

The corresponding perf-defconfig reference is declared in **apq8009.conf** file as shown below. The corresponding **apq8009.conf** reference file may be found under the following paths :- **“/poky/meta-qt-bsp/conf/machine/apq8009.conf”**

```
KERNEL_DEFCONFIG = “apq8909_defconfig”
```

```
KERNEL_PERF_DEFCONFIG = “apq8909-perf_defconfig”
```

How to get the kernel config from device

```
adb pull /proc/config.gz
```

How to extract Kernel Config when pulling /proc/config.gz doesn't work

Please get the vmlinux and copy it to kernel source folder.

Run the below script from any kernel source folder on the vmlinux file.

```
./kernel/scripts/extract-ikconfig vmlinux > kernel_config.txt
```

Sometimes extracting of kernel_config file from vmlinux may not work, please follow below procedure to get .config file from source build Please check for the file as show below

```
ls -la ./out/target/product/msm<chipset>/obj/KERNEL_OBJ/.config
```

the file is a hidden file ".config" <.dot config>

NOTE: -

To get the kernel config from the device and extract the kernel config from vmlinux please enable the below CONFIG in defconfig.

```
CONFIG_IKCONFIG_PROC=Y
```

```
CONFIG_IKCONFIG=Y
```

How to disable configs

We can disable these configs by making y to n or we can just delete the config from perf-defconfig

Ex:- For MSM8953 go to /kernel/arch/arm64/configs/msmcortex-perf_defconfig

Disable the CONFIG_CGROUP_DEBUG config :-

```
+CONFIG_CGROUP_DEBUG=n
```

Or

```
- CONFIG_CGROUP_DEBUG=y
```

How to Review the kernel config

Pull the kernel config from the device using “adb pull /proc/config.gz” and review the kernel config whether changes are reflected or not.

If you have made changes in defconfig to enable/disable the CONFIG you can see the changes as below in defconfig.

Ex: -

```
CONFIG_IKCONFIG=y
```

-If you have enabled this CONFIG

CONFIG_IKCONFIG is not set

-If you have disabled this config

Please make sure below config should be disabled in perf-def config

CONFIG_SERIAL_CORE=y

- Config used to initialize serial driver for early printk.

CONFIG_SERIAL_CORE_CONSOLE=y

- Config used for UART console write – write a console message to a serial port.

CONFIG_SERIAL_MSM_HS=y

CONFIG_SERIAL_MSM_HSL=y

CONFIG_SERIAL_MSM_HSL_CONSOLE=y

- By enabling these Configs we can enable its onboard high speed serial port for machine based on MSM family of SoC's.

NOTE:-

By disabling SERIAL Configs we can improve the boot time and performance.

CONFIG_CGROUP_DEBUG=y

- This config enables a simple controller that exports debugging information about the cgroup (Control Group) framework.
- Please disable this config for performance.

CONFIG_SLUB_DEBUG=y

CONFIG_SLUB_DEBUG_ON=y

- SLUB has extensive debug support features. Disabling these can result in significant savings in code size. This also disables SLUB sysfs support. /sys/slab will not exist and there will be no support for cache validation etc.

CONFIG_MSM_SMD_DEBUG=y

- Disable this DEBUG config for performance.

CONFIG_DYNAMIC_DEBUG=y

- This config enable the dynamic printk() support.
- Dynamic debugging is controlled via the 'dynamic_debug/control' file, which is contained in the 'debugfs' filesystem. Thus, the debugfs filesystem must first be mounted before making use of this feature.

-

CONFIG_DEBUG_PAGEALLOC=y

- Unmap pages from the kernel linear mapping after free_pages(). Helps to find certain types of memory corruptions.
- This results in a large slowdown so Please Disable this config.

CONFIG_DEBUG_KMEMLEAK=y

- If set 'Y', This Config enables the kernel memory leak detector.
- Enabling this feature will introduce an overhead to memory allocations. See Documentation/dev-tools/kmemleak.rst for more details.

CONFIG_DEBUG_KMEMLEAK_EARLY_LOG_SIZE=400

- Used for maximum kmemleak early log entries. If CONFIG_DEBUG_KMEMLEAK is not set this config not required to be set.

CONFIG_DEBUG_KMEMLEAK_DEFAULT_OFF=y

- Default kmemleak to off.

CONFIG_DEBUG_SPINLOCK=y

- Used for Debug spinlock usage.
- Disable this config for performance.

CONFIG_DEBUG_MUTEXES=y

- Used for Mutex debugging, deadlock detection.

CONFIG_DEBUG_ATOMIC_SLEEP=y

- Sleep inside atomic section checking.
- Various routines which may sleep will become very noisy if they are called inside atomic sections: when a spinlock is held, inside an rcu read side critical section, inside preempt disabled sections, inside an interrupt, etc..

CONFIG_DEBUG_STACK_USAGE=y

- Stack utilization instrumentation.

CONFIG_DEBUG_LIST=y

- Debug linked list manipulation.

CONFIG_LOCKUP_DETECTOR=y

- Detect Hard and soft Lockups.

CONFIG_SCHED_DEBUG=y

- Collect scheduler debugging info, /proc/sched_debug file will be provided for debug.

CONFIG_SYSRQ_SCHED_DEBUG=y

- CONFIG_SYSRQ_SCHED_DEBUG results in verbose printk output when the "show-task-states(T)" or "show-blocked-tasks(W)" sysrq triggers are invoked. Because this can increase the chances of a watchdog bark when the system is under heavy load, disable these debug prints.

CONFIG_DEBUG_MEMORY_INIT=y

- This config enables additional check during memory initialization.

CONFIG_DEBUG_PREEMPT=y

- With this config enable kernel will use a debug variant of the commonly used `smp_processor_id()` function and will print warnings if kernel code uses it in a preemption-unsafe way. Also, the kernel will detect preemption count underflows.

CONFIG_PAGE_POISONING=y

- Poison page after freeing
- Fill the pages with poison patterns after `free_pages()` and verify the patterns before `alloc_pages()`. This results in a large slowdown, but helps to find certain types of memory corruption.
- This option cannot be enabled in combination with hibernation as that would result in incorrect warnings of memory corruption after a resume because free pages are not saved to the suspend image.

CONFIG_RMNET_DATA_DEBUG_PKT=y

- Enable Debug for RMNET data to support MAP data feature for data connectivity.

CONFIG_BLK_DEV_LOOP=y

- Loopback device support.

CONFIG_RCU_CPU_STALL_VERBOSE=y

- Enable additional checks to print additional per-task information for `RCU_CPU_STALL_DETECTOR`.
- This option causes RCU to printk detailed per-task information for any tasks that are stalling the current RCU grace period.

CONFIG_TRACEPOINTS=y

- Enables the kernel trace points for debugging.

CONFIG_DEBUG_BUGVERBOSE=y

- Verbose `BUG()` Reporting.
- If set Y, make `BUG()` panics output the file name and line number of the `BUG` call as well as the EIP and oops trace. This aids debugging but costs about 70-100K of memory.

CONFIG_CC_OPTIMIZE_FOR_SIZE=y

- By disabling this config, we can get boot time improvement ~100ms.
- This config caused gcc to generate jump-to-jump code which causes cache line bouncing, hurting performance. So disable this config.

CONFIG_DEBUG_SET_MODULE_RONX=y

- Set loadable kernel module data as NX(No-Execute) and text as RO(Read-only).

CONFIG_CORESIGHT=y

CONFIG_CORESIGHT_EVENT=y

CONFIG_CORESIGHT_FUSE=y

CONFIG_CORESIGHT_CTI=y

CONFIG_CORESIGHT_TMC=y

CONFIG_CORESIGHT_TPIU=y

CONFIG_CORESIGHT_FUNNEL=y

CONFIG_CORESIGHT_REPLICATOR=y

CONFIG_CORESIGHT_STM=y

CONFIG_CORESIGHT_HWEVENT=y

CONFIG_CORESIGHT_ETMV4=y

CONFIG_CORESIGHT_MODEM_ETM=y

CONFIG_CORESIGHT_WCN_ETM=y

CONFIG_CORESIGHT_RPM_ETM=y

- All CORESIGHT Config used for CoreSight Tracing support.
- Disable All CORESIGHT Configs for performance.

CONFIG_RELAY=y-

- User space relay support
- This option enables support for relay interface support in certain file systems (such as debugfs).

CONFIG_FAULT_INJECTION=y

- Provide fault-injection framework. For more details, see Documentation/fault-injection/.

CONFIG_FAULT_INJECTION_STACKTRACE_FILTER=y

- Provide stacktrace filter for fault-injection capabilities

CONFIG_ALLOC_BUFFERS_IN_4K_CHUNKS=y

- Allocate buffer in page order 0, If we consider the case in which the size ≥ 64 KB
- Need to disable in case of page order ≥ 4 .
-

CONFIG_IOMMU_DEBUG=y

- Enable IOMMU Debugging.

- Force the IOMMU to on even when you have less than 4GB of memory and add debugging code. On overflow always panic. And allow to enable IOMMU leak tracing.
- When you use it make sure you have a big enough IOMMU/AGP aperture

Note: Compilation error while disabling CONFIG_CC_OPTIMIZE_FOR_SIZE=y

To disable CONFIG_CC_OPTIMIZE_FOR_SIZE=y Config requires some changes in main Makefile for compilation

diff --git a/Makefile b/Makefile

index 4045f28..326ca08 100644

--- a/Makefile

+++ b/Makefile

@@ -619,13 +619,15 @@ KBUILD_CFLAGS += \$(call cc-disable-warning,frame-address,)

KBUILD_CFLAGS += \$(call cc-disable-warning, format-truncation)

KBUILD_CFLAGS += \$(call cc-disable-warning, format-overflow)

KBUILD_CFLAGS += \$(call cc-disable-warning, int-in-bool-context)

+KBUILD_CFLAGS += \$(call cc-disable-warning, array-bounds)

KBUILD_CFLAGS += \$(call cc-option,-fno-PIE)

KBUILD_AFLAGS += \$(call cc-option,-fno-PIE)

ifdef CONFIG_CC_OPTIMIZE_FOR_SIZE

KBUILD_CFLAGS += -Os \$(call cc-disable-warning,maybe-uninitialized,)

else

-KBUILD_CFLAGS += -O2

+KBUILD_CFLAGS += -O2 \$(call cc-disable-warning,maybe-uninitialized,)

+

Endif

diff --git a/arch/arm/configs/msmcortex-perf_defconfig b/arch/arm/configs/msmcortex-perf_defconfig

index 1d28f37..264f22c 100644


```
--- a/arch/arm/configs/msmcortex-perf_defconfig
+++ b/arch/arm/configs/msmcortex-perf_defconfig
@@ -27,7 +27,6 @@ CONFIG_NAMESPACES=y

CONFIG_BLK_DEV_INITRD=y

CONFIG_RD_BZIP2=y

CONFIG_RD_LZMA=y
-CONFIG_CC_OPTIMIZE_FOR_SIZE=y

CONFIG_KALLSYMS_ALL=y

CONFIG_EMBEDDED=y

# CONFIG_SLUB_DEBUG is not set
```

Qualcomm

2018-05-02 23:44:52 PDT
zhangxiaochun2@huaiqin.com