

POWER GATING

Power Gating is effective for reducing leakage power [3]. Power gating is the technique wherein circuit blocks that are not in use are temporarily turned off to reduce the overall leakage power of the chip. This temporary shutdown time can also call as "low power mode" or "inactive mode". When circuit blocks are required for operation once again they are activated to "active mode". These two modes are switched at the appropriate time and in the suitable manner to maximize power performance while minimizing impact to performance. Thus goal of power gating is to minimize leakage power by temporarily cutting power off to selective blocks that are not required in that mode.

Power gating affects design architecture more compared to the clock gating. It increases time delays as power gated modes have to be safely entered and exited. The possible amount of leakage power saving in such low power mode and the energy dissipation to enter and exit such mode introduces some architectural trade-offs. Shutting down the blocks can be accomplished either by software or hardware. Driver software can schedule the power down operations. Hardware timers can be utilized. A dedicated power management controller is the other option.

An externally switched power supply is very basic form of power gating to achieve long term leakage power reduction. To shutoff the block for small interval of time internal power gating is suitable. CMOS switches that provide power to the circuitry are controlled by power gating controllers. Output of the power gated block discharge slowly. Hence output voltage levels spend more time in threshold voltage level. This can lead to larger short circuit current.

Power gating uses low-leakage PMOS transistors as header switches to shut off power supplies to parts of a design in standby or sleep mode. NMOS footer switches can also be used as sleep transistors. Inserting the sleep transistors splits the chip's power network into a permanent power network connected to the power supply and a virtual power network that drives the cells and can be turned off.

The quality of this complex power network is critical to the success of a power-gating design. Two of the most critical parameters are the IR-drop and the penalties in silicon area and routing resources. Power gating can be implemented using cell- or cluster-based (or fine grain) approaches or a distributed coarse-grained approach.

Power-gating parameters

Power gating implementation has additional considerations than the normal timing closure implementation. The following parameters need to be considered and their values carefully chosen for a successful implementation of this methodology [1] [2].

- **Power gate size:** The power gate size must be selected to handle the amount of switching current at any given time. The gate must be bigger such that there is no measurable voltage (IR) drop due to the gate. Generally we use 3X the switching capacitance for the gate size as a rule of thumb. Designers can also choose between header (P-MOS) or footer (N-MOS) gate. Usually footer gates tend to be smaller in area for the same switching current. Dynamic power analysis tools can accurately measure the switching current and also predict the size for the power gate.
- **Gate control slew rate:** In power gating, this is an important parameter that determines the power gating efficiency. When the slew rate is large, it takes more time to switch off and switch-on the circuit and hence can affect the power gating efficiency. Slew rate is controlled through buffering the gate control signal.
- **Simultaneous switching capacitance:** This important constraint refers to the amount of circuit that can be switched simultaneously without affecting the power network integrity. If a large amount of the circuit is switched simultaneously, the resulting "rush current" can compromise the power network integrity. The circuit needs to be switched in stages in order to prevent this.
- **Power gate leakage:** Since power gates are made of active transistors, leakage is an important consideration to maximize power savings.

Fine-grain power gating

Adding a sleep transistor to every cell that is to be turned off imposes a large area penalty, and individually gating the power of every cluster of cells creates timing issues introduced by inter-cluster voltage variation that are difficult to resolve. Fine-grain power gating encapsulates the switching transistor as a part of the standard cell logic. Switching transistors are designed by either library IP vendor or standard cell designer. Usually these cell designs conform to the normal standard cell rules and can easily be handled by EDA tools for implementation.

The size of the gate control is designed with the worst case consideration that this circuit will switch during every clock cycle resulting in a huge area impact. Some of the recent designs implement the fine-grain power gating selectively, but only for the low V_t cells. If the technology allows multiple V_t libraries, the use of low V_t devices is minimum in the design (20%), so that the area impact can be reduced. When using power gates on the low V_t cells the output must be isolated if the next stage is a high V_t cell. Otherwise it can cause the neighboring high V_t cell to have leakage when output goes to an unknown state due to power gating.

Gate control slew rate constraint is achieved by having a buffer distribution tree for the control signals. The buffers must be chosen from a set of always on buffers (buffers without the gate control signal) designed with high V_t cells. The inherent difference between when a cell switches off with respect to another, minimizes the rush current during switch-on and switch-off.

Usually the gating transistor is designed as a high v_t device. Coarse-grain power gating offers further flexibility by optimizing the power gating cells where there is low switching activity. Leakage optimization has to be done at the coarse grain level, swapping the low leakage cell for the high leakage one. Fine-grain power gating is an elegant methodology resulting in up to 10X leakage reduction. This type of power reduction makes it an appealing technique if the power reduction requirement is not satisfied by multiple V_t optimization alone.

Coarse-grain power gating

The coarse-grained approach implements the grid style sleep transistors which drives cells locally through shared virtual power networks. This approach is less sensitive to PVT variation, introduces less IR-drop variation, and imposes a smaller area overhead than the cell- or cluster-based implementations. In coarse-grain power gating, the power-gating transistor is a part of the power distribution network rather than the standard cell.

There are two ways of implementing a coarse-grain structure:

- 1) Ring-based
- 2) column-based

- **Ring-based methodology:** The power gates are placed around the perimeter of the module that is being switched-off as a ring. Special corner cells are used to turn the power signals around the corners.
- **Column-based methodology:** The power gates are inserted within the module with the cells abutted to each other in the form of columns. The global power is the higher layers of metal, while the switched power is in the lower layers.

Gate sizing depends on the overall switching current of the module at any given time. Since only a fraction of circuits switch at any point of time, power gate sizes are smaller as compared to the fine-grain switches. Dynamic power simulation using worst case vectors can determine the worst case switching for the module and hence the size. IR drop can also be factored into the analysis. Simultaneous switching capacitance is a major consideration in coarse-grain power gating implementation. In order to limit simultaneous switching daisy chaining the gate control buffers, special counters are used to selectively turn on blocks of switches.

Isolation Cells

Isolation cells are used to prevent short circuit current. As the name indicates these cells isolate power gated block from the normally on block. Isolation cells are specially designed for low short circuit current when input is at threshold voltage level. Isolation control signals are provided by power gating controller. Isolation of the signals of a switchable module is essential to preserve design integrity. Usually a simple OR or AND logic can function as an output isolation device. Multiple state retention schemes are available in practice to preserve the state before a module shuts down. The simplest technique is to scan out the register values into a memory before shutting down a module. When the module wakes up, the values are scanned back from the memory.

Retention Registers

When power gating is used, the system needs some form of state retention, such as scanning out data to a RAM, then scanning it back in when the system is reawakened. For critical applications, the memory states must be maintained within the cell, a condition that requires a retention flop to store bits in a table. That makes it possible to restore the bits very quickly during wakeup. Retention registers are special low leakage flip-flops used to hold the data of main register of the power gated block. Thus internal state of the block

during power down mode can be retained and loaded back to it when the block is reactivated. Retention registers are always powered up. The retention strategy is design dependent. During the power gating data can be retained and transferred back to block when power gating is withdrawn. Power gating controller controls the retention mechanism such as when to save the current contents of the power gating block and when to restore it back

Source : <http://asic-soc.blogspot.in/2008/04/power-gating.html>