Paper Review of

**A LANDSCAPE OF THE NEW DARK SILICON DESIGN REGIME**

Name: Abhishek S Yellanki

SBU ID: 111095603

**Q1.** Due to power constraints, only some part of the chip can function at full frequency and maximize efficiency. The left over part, which remains inactive for that period of time is termed “dark silicon”. This part of the processor cannot be used because of power constraints. If the whole chip has to be used it has to run at a lower frequency so that the chip maintains its power budget without overheating.

Due to this chip’s, efficiency increases only by 1.4x every two years which differs from the expected 2.8x by quite a huge margin and as a result the amount of dark silicon on the chip increases per process generation.

From a statistical point of view, the dark silicon area on the chip will be 93.75% after 8 years making only 7.25% of the chip useable.

**Q2.** According to Moore’s law the number of transistors that can be placed on a chip double every two years keeping the power constraints constant. The frequency of operation remains same. The performance gain should be exponential as predicted by Moore’s law. But when the power constraints are taken into consideration, the gains must be offset by transistor energy efficiency.

Initially when the transistor number increased, power remained constant as voltage and current scaled downwards and leakage currents were negligible. As the number of transistors on the chip increased, the leakages started to show considerable effect on the operating voltage of the chip which resulted in increased power consumption. To keep the power utilization constant, only some part of the chip is active and usable at full capacity. This power utilization barrier is creating the utilization wall due to which dark silicon area comes into picture.

In table 1. The operating voltage is not offset in the post-Dennard period which results in exponential increase in power consumption of the chip. This is causing the power constraint to break due to which the chip could be damaged. The Utilization wall is placed to keep the power constraint as per budget. Due to this dark silicon on the chip is increasing exponentially

**Q3.** Figure 1 talks about how dark silicon increases with advancement in chip design. As the number of transistors increases, power consumption increases which is not feasible. It can be traded off by keeping the frequency lower or keeping inactive part on the chip.

On the top design, multiple cores are used keeping the frequency constant. On the bottom design the frequency is increased whereas number of cores remains same. On the top design the dark silicon percentage decreases but instead the chip is utilized to perform at a lower rate. On the bottom design, the chip functions at full speed but at any given time, there is relatively more unused chip.

**Q4.**

1. **Shrinking horseman**: this approach talks about decreasing the chip size so that there is lesser dark silicon area on each chip. The cost of manufacturing of the chip decreases but other costs like chip design, I/O pad area, mask costs won’t go down. Moreover, if area decreases peak hotspot temperature rises due to which energy –efficiency gain is less than 40% which was the case when the chip was not shrunk.
2. **The dim horseman:**  In this approach instead of reducing the chip size, the dark silicon will be utilized by running the processor at lower frequencies or by utilizing it less frequently so that it doesn’t remain unused. The design can temporarily be made to run at higher frequencies which might increase temperature and hitting the thermal capacitance limits of the circuit, then under-clocking the chip so that normalcy is returned. Turbo-boost, computational sprinting, near threshold voltage (NTV) are some of the techniques to make use of the dark silicon.
3. **The Specialized horseman:** In this approach, the dark silicon area is specially assigned to perform dedicated functions. Clock gating is a technique used to turn on a part of the chip temporarily to perform any of the dedicated functions and then turn if off when switching to general purpose logic. In this case the simple lines of communication between the programmer and the software and the hardware break as different parts of the chip need different design features.
4. **The ‘deus ex machina’ horseman:** In this approach instead of relying on the existing fundamental logic gate, researching on a new type of logic gate that performs better in terms of power, than the present one is given priority so that the dark silicon component can be reduced.

I think the “specialized horseman” is better compared to other approaches described in the paper. Although there are cons for this approach, the valuable chip area is used in its entirety rather than sitting idle. Parallel processing can also be increased to its limit when this approach is used. This increase the efficiency of the system more than anticipated.

The least promising is the Shrinking horseman approach as the efficiency of the chip increases lessthan that compared to utilizing the chip with dark silicon area on it.