EEE4084F 2014
Quiz 1

Lectures 1-4 + Seminar 1

Question 1:

(a) Composition

(b) No. According to the model, there needs to be exactly 2x 4GB i.e. 8GB of memory to operate correctly.

(c) This is an aggregation relation. It implies that zero or more programs can run on the O/S, but the O/S is still operational even if there is no program running.

(d) Nothing is said about the data. You could argue that at most 8GB of physical memory used by the program (code and data) is a limit.

(e) There are at least two inaccuracies:
   a. A PC is not usually contained inside the monitor – not usually anyway (although nowadays you can get a PC built into the display, slate PC type of thing).
   b. That a PC is limited to exactly 8GB of memory.
   c. That a PC has exactly 4 CPUs.

Question 2:

(a) HPEC = High Performance Embedded Computer.
The embedded aspect of HPEC captures the task-specific nature of certain types of high performance systems, and the need for special-purpose computing platforms to be design specifically for these particular types of application. These special-purpose computers cater for applications such as radar, autonomous flight, sophisticated control systems and a broad variety of other systems, many of which can be generalized as computing systems that need real-time or fast response, deal with 'big data' yet only need a comparatively small form factor (i.e., compared to e.g. traditional supercomputer / general purpose computer cluster).

(b) See table on next page (based on Lecture 2)

Aspect	Hardware	Reconfigurable	Software
Description:	This situation is where all the computation is done on a hardware solution, e.g. an ASIC or custom designed PCB. No processor or FPGA included.	This is in some ways a combination of both: a reconfigurable platform allows at least for its internal interconnects to be reconfigured, or more generally allows parts of its hardware to be dynamically changeable from software control.	This is a micro-processor based computer system (or a general purpose processing platform) that uses one or more CPUs.
Advantages:	High speed & performance. Efficient (possibly lower power used than an idle mciro proc.). Highly parallelizable. Much flexibility in terms of hardware to use.	Faster than software alone. More flexible than just option to change of software. More flexible than hardware because s/w and firmware aspects can be changed. Advantages of potential for supporting highly parallelizable solutions. i.e essentially has many of the advantages offered by both h/w and s/w.	Flexible Adaptable Can be much cheaper
Disadvantages:	Expensive. Static (cannot change). Difficulties in reuse / repurposing for other applications.	Expensive. Complex. Lengthy learning curve. May need specialized development tool chain to build applications. Essentially also has all the drawbacks of both h/w and s/w.	The hardware is static. Limits of clock speed and sequential processing.

Question 3:

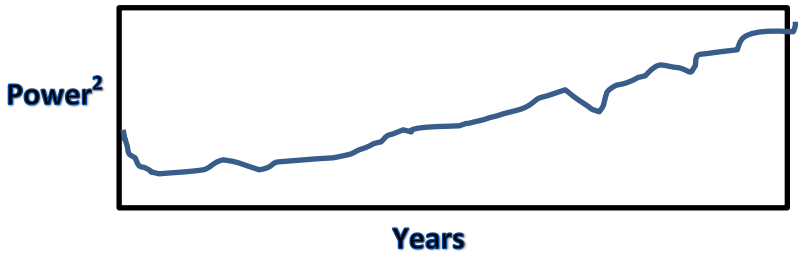
(a) 1000s of processors per chip tends to be a better fit for the classic and well-understood paradigm of working with standard processors that each execute a limited set of instructions. This approach provides a marked simplification over the situation in which programmers need to learn specialized instructions sets or gain a detailed understanding of particular platform design characteristics in order to develop an application. Furthermore, the approach of many simple processors provides significant benefits of code reuse. While supporting reuse the strategy also offers the potential for highly parallel solutions, allowing for scalability and clustering of the CPUs based on the granularity of the problem.

(b) A DWARF is a processing component that implements a computation design pattern and/or inter-process communication strategy. The approach is to have the DWARFs working together and when possible concurrently in order to provide a reliable parallel solution. Essentially, these DWARFs are well tested and reliable processes that perform a well-defined computation and offer a clear but flexible interface suited to reuse.

(c) The CW stands for “conventional wisdom”… The concept of CW relates to trends or design recommendations or tendencies that are considered appropriate and perhaps even recommendable for a particular era.
    a. Old CW = The conventional wisdom of the previous ‘era’ of computing, as in prior to year 2000. An example of an Old CW is “Power is free, but transistors are expensive”.
    b. New CW = The new wisdom suited to the current era of computing, as in post year 2000. A corresponding CW to the one above is the Power Wall: “Power is expensive, but transistors are ‘free’”.

Question 4:

(a) Power concerns: simply increasing (as in doubling) the transistors can lead to an exponential growth, likely more than an equivalent double, in terms of power draw. For instance, more transistors, and more parallel operations, can lead to more heat. This means more cooling. And cooling can be costly, indeed more costing than heating (which is likely why we were having a power cut in summer). In terms of running a high performance computing centre for instance, the is a lot of power draw in terms of the many computer and the collection of their individual power draw adds up to a significant amount, and is further made worse by the needs for cooling (e.g. running air conditioning to keep the centres at a cool). Thusly, and in line with the new CW concerning power use, it is important to ‘keep a lid’ on power demanded by individual processors otherwise the cost of running computer centers becomes exorbitant. Furthermore, restricting power draw clearly provides more wide-reaching benefits, such as in terms of reducing power draw on the grid, possibility for being able to run data centres on generators for period of time when the grid power is unavailable, and potentially allowing more scalability to data centres being able to maintain their existing power envelope as opposed to need grid upgrades. Consequently sensible limitations on powerdraw can have a far reaching impact such as improving the potential for running computers from renewable, but variable and inconsistent, power sources, such as wind or solar power. In terms of mobile computers, there are further benefits to working with power constraints, for instance aiming towards battery powered computers that can run for longer periods between recharging.
(b) This trend still holds for recent processors, in terms of the design going more parallel, although individual processor cores are not doubling in speed as in earlier years. For example P4, Dual core, quad core, etc., are using more and more transistors as per the Moore’s Law trend.
(c) Generally, pretty much (based on the lectures and much of what the text book has to say), but not as consistent. I would be expecting more an answer of Yes more than a definite No. Certainly, from around 70’s to around 2004, there has been something of an exponential growth in power draw, where computers every couple of years would be drawing something like a doubling the power every 18 months. But there are exceptions, some major ones for example the dip in power trend line between ENIAC and Intel 4004 for example were one to go back that far. But also definite exceptions after around 2004 where improvements and innovations in terms of power saving have been put in place – but as diagrams indicated in the lecture if the processor were to be running at full tilt then, yes, it’s pretty much an exponential power draw trend still. The figure below gives an impression.



Question 5:

There are many ways this question would be answered. Answers to this question will be marked according to how well the student has understood these abstract algorithmic and architectural constraints, and has provided a clear and suitably motivated explanation. Specific marks aren't allocated to the diagram, but this will contribute towards getting a better mark if it suitably aids the textual explanation.

Question 6: (1 bonus mark)

The answer is (i), as stated in the text: the Intel Paragon was based on the Intel i860 programmable microprocessor running at 50 MHz and performing at about 0.07 MFLOPS/W.