Operating systems II
CS 2506

Dr. Dan Grigoras

School of Computer Science and Information Technology

Computing systems

- What is a computer now? Hardware, software, but connectivity plays an ever-increasing role by giving access to remote resources.
- The computer architecture and configuration changed during time:
  - CPU/processor/core ➔ many cores, homogeneous or heterogeneous;
  - Main memory (plain) ➔ memory hierarchy (cache, main memory);
  - Storage (local or remote), disk ➔ array of large disks, SSD;
  - Bus ➔ multiple, faster buses;
  - Peripherals ➔ + Sensors
  - Stationary or mobile

- ...or a *cloud instance* provided as a service.

- *Question*: how are the hardware resources allocated to applications and managed?

Allocation of computing resources

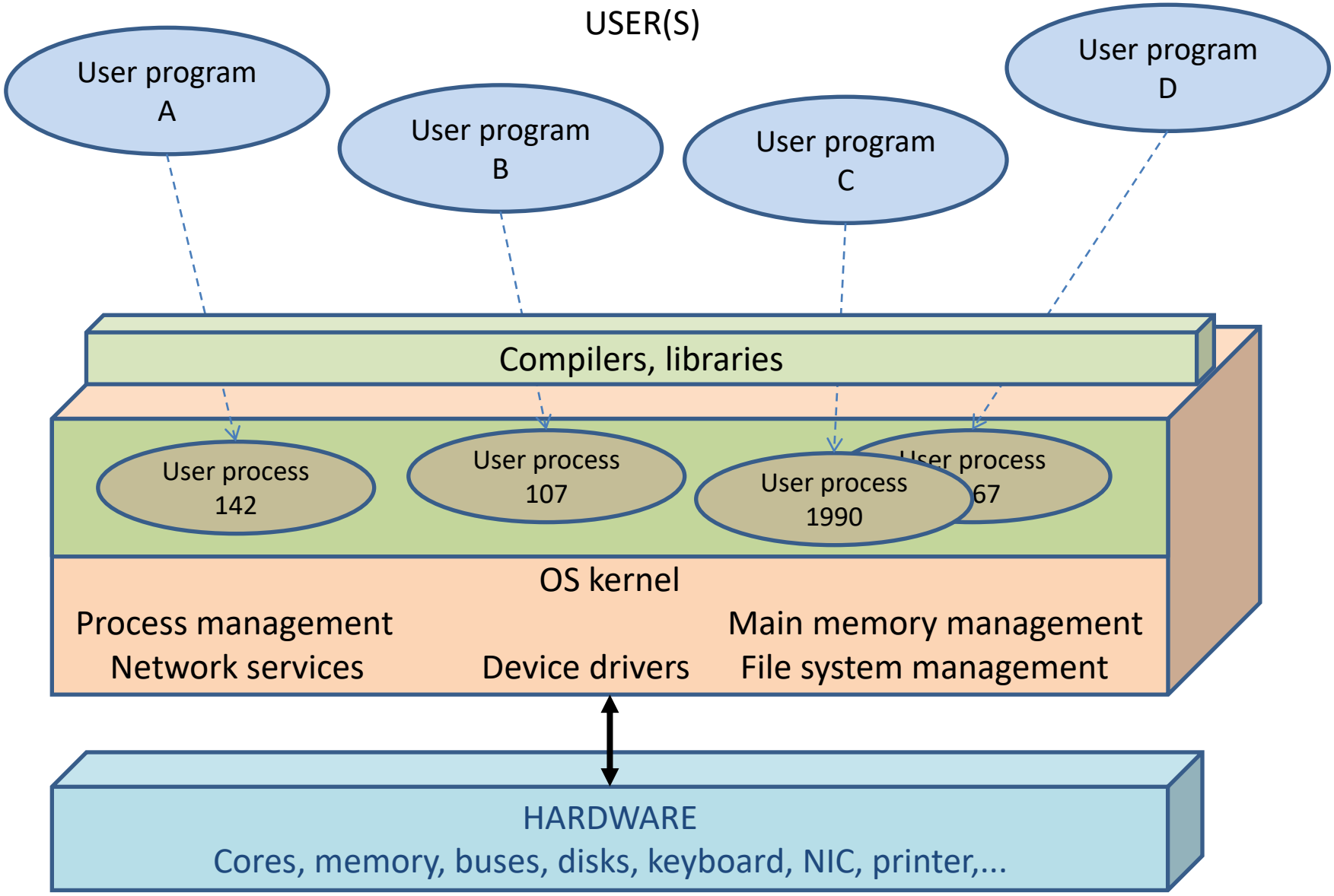
- The OS manages applications as *processes*. The OS purpose is to allocate the necessary resources for execution and to manage each process during its life cycle.
- Regarding resources, generally, there is more than one physical unit of a type of resource in a computer configuration.
  Examples: many cores, memory pages, disk sectors, networking cards, cameras.

- Resource allocation to processes is carried out in terms of units and is managed by specific OS services during processes life-time.

- Many operating systems are using virtualisation of resources such as cores and memory pages for more efficiency.

Computing resources management

- OS goals:
  - To meet user requirements
    - least execution time, quick response;
    - least cost;
    - fairness;
    - best possible user experience;
    - becurity.
  - To meet system admin requirements
    - optimal use of resources;
    - least energy use (green computing);
    - maximise revenue.
- The system admin can be the user/owner but not necessarily in all cases – see rack-scale and cloud computing.


USER(S)
User program A
User program B
User program C
User program D
Compilers, libraries
User process 142
User process 107
User process 1990
User process 67
OS kernel
Process management
Main memory management
Network services
Device drivers
File system management
HARDWARE
Cores, memory, buses, disks, keyboard, NIC, printer,...

Computing system layers

- The bottom layer is the hardware; it accepts basic commands such as "seek the disk arm to track 79, select head 3 and read sector 5". Software that interacts directly with the hardware is non-portable and contains low level details: data for control and state registers, interrupt priorities, DMA starting address,…

- The OS kernel has several key components:
  - *Device drivers* are the hardware units' managers; they hide the low level's details.
  - *Process management* handles processes, allocating them resources and scheduling their execution.
  - *Memory management* is responsible for physical memory and virtual memory management.
  - *File sys manager* is the code that organizes the data storage into files and directories, hiding low level details re disk blocks, for example.
  - *Network services* provide host-to-host and process-to-process communication across network.

Access to kernel services

- The kernel is a collection of services, some of which being accessible to user processes. They offer functionality and a higher level of abstraction of the computer. The repertoire of commands supported by the kernel defines the “virtual machine” which is platform-independent.
- To enter the kernel, the user process makes a *system call by executing a trap instruction of some sort.*
- This instruction switches the processor into a privileged operating mode (*kernel mode) and jumps into the kernel through a well-defined trap address.*
- Parameters passed with the trap instruction tell the kernel what service is requested.
- When the function is completed, the processor flips back to its normal operating mode (*user mode)* and *returns control to the user process.*

Trap instructions

- There are functions that require specific knowledge of handling resources – control registers, state register, sequence of operations, and a certain degree of protection – resources shared by several users/processes.
- These functions are coded as service routines; they are also known as *system calls*.

- The sequence of steps for a system call is:
  - The system call is invoked by the user process;
  - OS function is performed;
  - Control returns to the user process.

- Trap instructions are used to implement system calls.

Library functions

- User programs have access to libraries and can include in their code functions of these libraries – linked in the executables.
- Some library functions use system calls. The function itself parcels up the parameters correctly and then performs the trap.

- The function works as a wrapper for the system call.
- Example:

printf(“hello world”); ➔ write(1, “hello world”, 11);

OS classes

- There isn't one OS that fits all computing devices!
- The large range of computing devices requires customized OS.

- *General purpose computers* run general purpose OS: Unix, Linux, Windows,... that deal with different kind of processes, such as interactive, compute-bound or transaction-based processes.
- *Mobile devices*, such as smart phones or sensors, run OS such as iPhone OS, Android, TinyOS, concerned with power saving and user experience. Their processes are event-driven.
- *Application-oriented OS*: Robot OS, Home OS, Internetware

- *Embedded systems* (IoT) run scaled down versions of OS, real-time, event-driven.

Trends

- Rapid advances in hardware:
  - peripheral devices are integrated onto a die, each with complex, varying programming models, that execute specialized firmware with little OS integration.
  - for energy saving, cores, devices, and memory can be powered on/off at any time during the system operation.
  - large, non-volatile main memories will likely become prevalent and might have greater capacity than today's disks and disk arrays.
- Applications are changing:
  - many applications use cloud (data center) services (or microservices), individual users' mobile devices, sensors and actuators of the Internet of Things. Do we need an Internet OS?
  - Enterprise applications such as file servers, relational databases, and big data analytics now come pre-packaged in a *rack-scale software appliances*, delivered to customers who have only to plug in and use the system.
  - Clouds are running applications in containers or VMs with their own OS.

Suggested readings

1. H. Mei, Y. Gao: *Operating Systems for Internetware: Challenges and Future Directions*, IEEE 38th International Conference on Distributed Computing Systems (ICDCS) 2018 https://ieeexplore.ieee.org/document/8416401
2. D. Milojicic and T. Roscoe: *Will OSs in 2025 still resemble the Unix-like consensus of today, or will a very different design achieve widespread adoption?*, Computer, Volume: 49, Jan. 2016, pp. 43-51, http://ieeexplore.ieee.org.ucc.idm.oclc.org/document/7383138

- My comments: good overview of existing operating systems and technology advances that will have a great impact on operating systems architecture and operation. These papers can suggest interesting avenues of further study and research.

Course goals and methodology

- Goals
  - understand and learn the main services of an operating system:
    - processes and threads management, including scheduling and load balancing, IPC;
    - cores, memory management;
    - device drivers;
    - file management system.
  - understand and learn features of different models of operating systems:
    - general purpose;
    - mobile;
    - sensor.

- Methodology
  - attending the lectures and engaging in discussions on lectures' topics.
  - carrying out the lab work.
  - using recommended references to learn more about specific topics.

- Most of the topics are comprehensively presented in the book by

A. Tanenbaum and H. Bos: Modern Operating Systems, Pearson, 4th edition, 2014.

Canvas CS 2506

- Grading
  - Continuous assessment: 20%
  - Written exam: 90 min 80%

- Lectures: 50 min + 5 min review & questions

- Office: G65, Western Gateway Bldg
- Email: grigoras@cs.ucc.ie, Subject: CS 2506
- Office hours: by appointment

Plagiarism

1. Plagiarism is presenting someone else's work as your own. It is a violation of UCC Policy and there are strict and severe penalties.
2. You must read and comply with the UCC Policy on Plagiarism www.ucc.ie/en/exams/procedures-regulations/
3. The Policy applies to *all* work submitted, including software.
4. You can expect that your work will be checked for evidence of plagiarism or collusion.
5. In some circumstances it may be acceptable to reuse a small amount of work by others, but *only* if you provide explicit acknowledgement and justification.
6. If in doubt ask your module lecturer *prior* to submission. Better safe than sorry!