PKU--LL203-进程线程

阶段一：进程管理

- Computer multitasking: https://en.wikipedia.org/wiki/Computer_multitasking
    - In computing, multitasking is a concept of performing multiple tasks (also known as processes) over a certain period of time by executing them concurrently. New tasks start and interrupt already started ones before they have reached completion, instead of executing the tasks sequentially so each started task needs to reach its end before a new one is started.
    - Task: https://en.wikipedia.org/wiki/Task_(computing)
        - In computing, a task is a unit of execution. In some operating systems, a task is synonymous with a process, and in others with a thread. In batch processing computer systems, a task is a unit of execution within a job.
    - Process: https://en.wikipedia.org/wiki/Process_(computing)
        - In computing, a process is an instance of a computer program that is being executed.
        - It contains the program code and its current activity.
        - Depending on the operating system (OS), a process may be made up of multiple threads of execution that execute instructions concurrently.
    - Light-weight process: https://en.wikipedia.org/wiki/Light-weight_process
        - In computer operating systems, a light-weight process (LWP) is a means of achieving multitasking.
        - In the traditional meaning of the term, as used in Unix System V and Solaris, a LWP runs in user space on top of a single kernel thread and shares its address space and system resources with other LWPs within the same process.
        - On Linux, user threads are implemented by allowing certain processes to share resources, which sometimes leads to these processes to be called "light weight processes".
    - Thread: https://en.wikipedia.org/wiki/Thread_(computing)
        - In computer science, a thread of execution is the smallest sequence of programmed instructions that can be managed independently by a scheduler, which is typically a part of the operating system.
        - The implementation of threads and processes differs between operating systems, but in most cases a thread is a component of a process.
        - Multiple threads can exist within the same process, executing concurrently (one starting before others finish) and share resources such as memory, while different processes do not share these resources.
        - In particular, the threads of a process share its instructions (executable code) and its context (the values of its variables at any given moment).
- Process management: https://en.wikipedia.org/wiki/Process_management_(computing)
    - Process management is an integral part of any modern-day operating system (OS).
    - The OS must allocate resources to processes, enable processes to share and exchange information, protect the resources of each process from other processes and enable synchronisation among processes.
    - To meet these requirements, the OS must maintain a data structure for each process, which describes the state and resource ownership of that process, and which enables the OS to exert control over each process.
    - Process identifier: https://en.wikipedia.org/wiki/Process_identifier
        - In computing, the process identifier (normally referred to as the process ID or just PID) is a number used by most operating system kernels,—such as that of UNIX, Mac OS X or Microsoft Windows—to uniquely identify an active process.
        - This number may be used as a parameter in various function calls allowing processes to be manipulated, such as adjusting the process's priority or killing it altogether.
    - Process control block: https://en.wikipedia.org/wiki/Process_control_block
        - Process Control Block (PCB, also called Task Controlling Block, process table, Task Struct, or Switchframe) is a data structure in the operating system kernel containing the information needed to manage a particular process.
    - Thread control block: https://en.wikipedia.org/wiki/Thread_control_block
        - Thread Control Block (TCB) is a data structure in the operating system kernel which contains thread-specific information needed to manage it.
- Fork-exec: https://en.wikipedia.org/wiki/Fork-exec
    - Fork-exec is a commonly used technique in Unix whereby an executing process spawns a new program.
    - Parent process: https://en.wikipedia.org/wiki/Parent_process
        - In computing, a parent process is a process that has created one or more child processes.
    - Child process: https://en.wikipedia.org/wiki/Child_process
        - A child process in computing is a process created by another process (the parent process). This technique pertains to multitasking operating systems. There are two major procedures for creating a child process: the fork system call (preferred in Unix-like systems and the POSIX standard) and the spawn (preferred in the modern (NT) kernel of Microsoft Windows, as well as in some historical operating systems).
    - Orphan process: https://en.wikipedia.org/wiki/Orphan_process
        - An orphan process is a computer process whose parent process has finished or terminated, though it remains running itself.
    - Process group: https://en.wikipedia.org/wiki/Process_group

        - In POSIX-conformant operating systems, a process group denotes a collection of one or more processes. Process groups are used to control the distribution of signals. A signal directed to a process group is delivered individually to all of the processes that are members of the group.
        - Process groups are themselves grouped into sessions.
    - Thread pool pattern: https://en.wikipedia.org/wiki/Thread_pool_pattern
        - In computer programming, the thread pool pattern (also replicated workers or worker-crew model) is where a number of threads are created to perform a number of tasks, which are usually organized in a queue.
    - Fork-join model: https://en.wikipedia.org/wiki/Fork%E2%80%93join_model
        - In parallel computing, the fork–join model is a way of setting up and executing parallel programs, such that execution branches off in parallel at designated points in the program, to "join" (merge) at a subsequent point and resume sequential execution.
    - pstree: https://en.wikipedia.org/wiki/Pstree
        - 执行：pstree命令
- Process state: https://en.wikipedia.org/wiki/Process_state
    - The following typical process states are possible on computer systems of all kinds. In most of these states, processes are "stored" on main memory.
        - Created
        - Ready and waiting
        - Running
        - Blocked
        - Terminated
    - Run queue: https://en.wikipedia.org/wiki/Run_queue
        - In modern computers many processes run at once. Active processes are placed in an array called a run queue, or runqueue. The run queue may contain priority values for each process, which will be used by the scheduler to determine which process to run next.
    - Blocking: https://en.wikipedia.org/wiki/Blocking_(computing)
        - In computing, a process is an instance of a computer program that is being executed. A process always exists in exactly one process state. A process that is blocked is one that is waiting for some event, such as a resource becoming available or the completion of an I/O operation.
    - Zombie process: https://en.wikipedia.org/wiki/Zombie_process
        - On Unix and Unix-like computer operating systems, a zombie process or defunct process is a process that has completed execution (via the exit system call) but still has an entry in the process table: it is a process in the "Terminated state". This occurs for child processes, where the entry is still needed to allow the parent process to read its child's exit status: once the exit status is read via the wait system call, the zombie's entry is removed from the process table and it is said to be "reaped".

自学：

1. Computer multitasking: https://en.wikipedia.org/wiki/Computer_multitasking
2. Task: https://en.wikipedia.org/wiki/Task_(computing)
3. Process: https://en.wikipedia.org/wiki/Process_(computing)
4. Light-weight process: https://en.wikipedia.org/wiki/Light-weight_process
5. Thread: https://en.wikipedia.org/wiki/Thread_(computing)
6. Process management: https://en.wikipedia.org/wiki/Process_management_(computing)
7. Process identifier: https://en.wikipedia.org/wiki/Process_identifier
8. Process control block: https://en.wikipedia.org/wiki/Process_control_block
9. Thread control block: https://en.wikipedia.org/wiki/Thread_control_block
10. Fork-exec: https://en.wikipedia.org/wiki/Fork-exec
11. Parent process: https://en.wikipedia.org/wiki/Parent_process
12. Child process: https://en.wikipedia.org/wiki/Child_process
13. Orphan process: https://en.wikipedia.org/wiki/Orphan_process
14. Process group: https://en.wikipedia.org/wiki/Process_group
15. Thread pool pattern: https://en.wikipedia.org/wiki/Thread_pool_pattern
16. Fork-join model: https://en.wikipedia.org/wiki/Fork%E2%80%93join_model
17. pstree: https://en.wikipedia.org/wiki/Pstree
18. Process state: https://en.wikipedia.org/wiki/Process_state
19. Run queue: https://en.wikipedia.org/wiki/Run_queue
20. Blocking: https://en.wikipedia.org/wiki/Blocking_(computing)
21. Zombie process: https://en.wikipedia.org/wiki/Zombie_process

---

阶段二：进程调度

- Time-sharing: https://en.wikipedia.org/wiki/Time-sharing
    - In computing, time-sharing is the sharing of a computing resource among many users by means of multiprogramming and multi-tasking. Its introduction in the 1960s by students and professors at Dartmouth College, and emergence as the prominent model of computing in the 1970s, represented a major technological shift in the history of computing.

  - Concurrency: https://en.wikipedia.org/wiki/Concurrency_(computer_science)
    - In computer science, concurrency is a property of systems in which several computations are executing simultaneously, and potentially interacting with each other.
    - The computations may be executing on multiple cores in the same chip, preemptively time-shared threads on the same processor, or executed on physically separated processors.
    - Concurren computing: https://en.wikipedia.org/wiki/Concurrent_computing
      - Concurrent computing is a form of computing in which several computations are executing during overlapping time periods – concurrently – instead of sequentially (one completing before the next starts). This is a property of a system – this may be an individual program, a computer, or a network – and there is a separate execution point or "thread of control" for each computation ("process").
    - Concurrency control: https://en.wikipedia.org/wiki/Concurrency_control
      - In information technology and computer science, especially in the fields of computer programming, operating systems, multiprocessors, and databases, concurrency control ensures that correct results for concurrent operations are generated, while getting those results as quickly as possible.
- Scheduling: https://en.wikipedia.org/wiki/Scheduling_(computing)
  - In computing, scheduling is the method by which work specified by some means is assigned to resources that complete the work.
  - The resources may be virtual computation elements such as threads, processes or data flows, which are in turn scheduled onto hardware resources such as processors, network links or expansion cards.
  - Preemption: https://en.wikipedia.org/wiki/Preemption_(computing)
    - In computing, preemption is the act of temporarily interrupting a task being carried out by a computer system, without requiring its cooperation, and with the intention of resuming the task at a later time.
    - Such a change is known as a context switch. It is normally carried out by a privileged task or part of the system known as a preemptive scheduler, which has the power to preempt, or interrupt, and later resume, other tasks in the system.
  - Non-preemptive multitasking: https://en.wikipedia.org/wiki/Non-preemptive_multitasking
    - Non-preemptive multitasking or cooperative multitasking is a style of computer multitasking in which the operating system never initiates a context switch from a running process to another process. Instead, processes voluntarily stop their operation periodically in order to enable multiple applications to be run simultaneously.
    - This type of multitasking is called "cooperative" because all programs must cooperate for the entire scheduling scheme to work. In this scheme, the process scheduler of an operating system is known as cooperative scheduler, having its role reduced down to starting the processes and letting them return control back to it voluntarily.
  - CPU-bound: https://en.wikipedia.org/wiki/CPU-bound
    - In computer science, a computer is CPU-bound (or compute-bound) when the time for it to complete a task is determined principally by the speed of the central processor: processor utilization is high, perhaps at 100% usage for many seconds or minutes. Interrupts generated by peripherals may be processed slowly, or indefinitely delayed.
  - I/O bound: https://en.wikipedia.org/wiki/I/O_bound
    - In computer science, I/O bound refers to a condition in which the time it takes to complete a computation is determined principally by the period spent waiting for input/output operations to be completed. This is the opposite of a task being CPU bound. This circumstance arises when the rate at which data is requested is slower than the rate it is consumed or, in other words, more time is spent requesting data than processing it.
  - Scheduling policies in linux kernel: see Documentation/scheduler/ and kernel/sched/
    - include/uapi/linux/sched.h
    - SCHED_NORMAL
      - Completely fair scheduler: https://en.wikipedia.org/wiki/Completely_Fair_Scheduler
        - The Completely Fair Scheduler (CFS) is a process scheduler which was merged into the 2.6.23 release of the Linux kernel and is the default scheduler. It handles CPU resource allocation for executing processes, and aims to maximize overall CPU utilization while also maximizing interactive performance.
    - SCHED_FIFO
      - FIFO: https://en.wikipedia.org/wiki/FIFO_(computing_and_electronics)
        - FIFO is an acronym for First In, First Out, a method for organizing and manipulating a data buffer, where the oldest (first) entry, or 'head' of the queue, is processed first.
        - It is analogous to processing a queue with first-come, first-served (FCFS) behaviour: where the people leave the queue in the order in which they arrive.
        - FCFS is also the jargon term for the FIFO operating system scheduling algorithm, which gives every process CPU time in the order in which it is demanded.
    - SCHED_RR
      - Round-robin scheduling: https://en.wikipedia.org/wiki/Round-robin_scheduling
        - Round-robin (RR) is one of the algorithms employed by process and network schedulers in computing. As the term is generally used, time slices are assigned to each process in equal portions and in circular order, handling all processes without priority (also known as cyclic executive). Round-robin scheduling is simple, easy to implement, and starvation-free.

- Context switch: https://en.wikipedia.org/wiki/Context_switch
    - In computing, a context switch is the process of storing and restoring the state (more specifically, the execution context) of a process or thread so that execution can be resumed from the same point at a later time. This enables multiple processes to share a single CPU and is an essential feature of a multitasking operating system.
    - Execution: https://en.wikipedia.org/wiki/Execution_(computing)
        - Execution in computer and software engineering is the process by which a computer or a virtual machine performs the instructions of a computer program.
        - The instructions in the program trigger sequences of simple actions on the executing machine. Those actions produce effects according to the semantics of the instructions in the program.
    - Context: https://en.wikipedia.org/wiki/Context_(computing)
        - In computer science, a task context is the minimal set of data used by a task (which may be a process or thread) that must be saved to allow a task interruption at a given date, and a continuation of this task at the point it has been interrupted and at an arbitrary future date.
        - The concept of context assumes significance in the case of interruptible tasks, wherein upon being interrupted the processor saves the context and proceeds to serve the Interrupt service routine. Thus, the smaller the context is, the smaller the latency is.
    - Task state segment: https://en.wikipedia.org/wiki/Task_state_segment
        - The task state segment (TSS) is a special structure on x86-based computers which holds information about a task. It is used by the operating system kernel for task management.
        - Specifically, the following information is stored in the TSS:
            - Processor register state
            - I/O port permissions
            - Inner-level stack pointers
            - Previous TSS link

自学：

1. Time-sharing: https://en.wikipedia.org/wiki/Time-sharing
2. Concurrency: https://en.wikipedia.org/wiki/Concurrency_(computer_science)
3. Concurrent computing: https://en.wikipedia.org/wiki/Concurrent_computing
4. Concurrency control: https://en.wikipedia.org/wiki/Concurrency_control
5. Scheduling: https://en.wikipedia.org/wiki/Scheduling_(computing)
6. Preemption: https://en.wikipedia.org/wiki/Preemption_(computing)
7. Non-preemptive multitasking: https://en.wikipedia.org/wiki/Non-preemptive_multitasking
8. CPU-bound: https://en.wikipedia.org/wiki/CPU-bound
9. I/O bound: https://en.wikipedia.org/wiki/I/O_bound
10. Completely fair scheduler: https://en.wikipedia.org/wiki/Completely_Fair_Scheduler
11. FIFO: https://en.wikipedia.org/wiki/FIFO_(computing_and_electronics)
12. Round-robin scheduling: https://en.wikipedia.org/wiki/Round-robin_scheduling
13. Context switch: https://en.wikipedia.org/wiki/Context_switch
14. Execution: https://en.wikipedia.org/wiki/Execution_(computing)
15. Context: https://en.wikipedia.org/wiki/Context_(computing)
16. Task state segment: https://en.wikipedia.org/wiki/Task_state_segment

---

阶段三：进程地址空间

- ELF: https://en.wikipedia.org/wiki/Executable_and_Linkable_Format
    - In computing, the Executable and Linkable Format (ELF, formerly called Extensible Linking Format) is a common standard file format for executables, object code, shared libraries, and core dumps.
    - Executable: https://en.wikipedia.org/wiki/Executable
        - In computing, an executable file or executable program, or sometimes simply an executable, causes a computer "to perform indicated tasks according to encoded instructions," as opposed to a data file that must be parsed by a program to be meaningful.
    - Object file: https://en.wikipedia.org/wiki/Object_file
        - An object file is a file containing object code, meaning relocatable format machine code that is usually not directly executable. There are various formats for object files, and the same object code can be packaged in different object files.
    - Library: https://en.wikipedia.org/wiki/Library_(computing)
        - In computer science, a library is a collection of non-volatile resources used by computer programs, often to develop software. These may include configuration data, documentation, help data, message templates, pre-written code and subroutines, classes, values or type specifications.
    - Momory Layout:
        - Code segment: https://en.wikipedia.org/wiki/Code_segment
            - In computing, a code segment, also known as a text segment or simply as text, is a portion of an object file or the corresponding section of the program's virtual address space that contains executable instructions.
        - Data segment: https://en.wikipedia.org/wiki/Data_segment
            - In computing, a data segment (often denoted .data) is a portion of an object file or the

corresponding virtual address space of a program that contains initialized static variables, that is, global variables and static local variables.

    - .bss: https://en.wikipedia.org/wiki/.bss
      - In computer programming, the name .bss or bss is used by many compilers and linkers for a part of the data segment containing statically-allocated variables represented solely by zero-valued bits initially (i.e., when execution begins). It is often referred to as the "bss section" or "bss segment".
    - Stack: https://en.wikipedia.org/wiki/Stack_(abstract_data_type)
      - In computer science, a stack or LIFO (last in, first out) is an abstract data type that serves as a collection of elements, with two principal operations: push, which adds an element to the collection, and pop, which removes the last element that was added.
    - Heap: (redirect to Memory management)
      - The task of fulfilling an allocation request consists of locating a block of unused memory of sufficient size. Memory requests are satisfied by allocating portions from a large pool of memory called the heap or free store.
- Loader: https://en.wikipedia.org/wiki/Loader_(computing)
  - In computing, a loader is the part of an operating system that is responsible for loading programs and libraries. It is one of the essential stages in the process of starting a program, as it places programs into memory and prepares them for execution. Loading a program involves reading the contents of the executable file containing the program instructions into memory, and then carrying out other required preparatory tasks to prepare the executable for running. Once loading is complete, the operating system starts the program by passing control to the loaded program code.
  - Dynamic loading: https://en.wikipedia.org/wiki/Dynamic_loading
    - Dynamic loading is a mechanism by which a computer program can, at run time, load a library (or other binary) into memory, retrieve the addresses of functions and variables contained in the library, execute those functions or access those variables, and unload the library from memory.
  - Relocation: https://en.wikipedia.org/wiki/Relocation_(computing)
    - Relocation is the process of assigning load addresses to various parts of a program and adjusting the code and data in the program to reflect the assigned addresses. A linker usually performs relocation in conjunction with symbol resolution, the process of searching files and libraries to replace symbolic references or names of libraries with actual usable addresses in memory before running a program.
  - Position-independent code: https://en.wikipedia.org/wiki/Position-independent_code
    - In computing, position-independent code (PIC) or position-independent executable (PIE) is a body of machine code that, being placed somewhere in the primary memory, executes properly regardless of its absolute address.
  - Copy-on-write: https://en.wikipedia.org/wiki/Copy-on-write
    - Copy-on-write (sometimes referred to as "COW") is an optimization strategy used in computer programming. Copy-on-write stems from the understanding that when multiple separate tasks use initially identical copies of some information (i.e., data stored in computer memory or disk storage), treating it as local data that they may occasionally need to modify, then it is not necessary to immediately create separate copies of that information for each task.
  - Memory-mapped file: https://en.wikipedia.org/wiki/Memory-mapped_file
    - A memory-mapped file is a segment of virtual memory which has been assigned a direct byte-for-byte correlation with some portion of a file or file-like resource. This resource is typically a file that is physically present on-disk, but can also be a device, shared memory object, or other resource that the operating system can reference through a file descriptor. Once present, this correlation between the file and the memory space permits applications to treat the mapped portion as if it were primary memory.
    - mmap: https://en.wikipedia.org/wiki/Mmap
      - In computing, mmap(2) is a POSIX-compliant Unix system call that maps files or devices into memory. It is a method of memory-mapped file I/O.
      - File-backed mapping maps an area of the process's virtual memory to files; i.e. reading those areas of memory causes the file to be read. It is the default mapping type.
      - Anonymous mapping maps an area of the process's virtual memory not backed by any file. The contents are initialized to zero. In this respect an anonymous mapping is similar to malloc, and is used in some malloc(3) implementations for certain allocations. However, anonymous mappings are not part of the POSIX standard, though implemented by almost all operating systems, by the MAP_ANONYMOUS flag.

自学：

1. ELF: https://en.wikipedia.org/wiki/Executable_and_Linkable_Format
2. Executable: https://en.wikipedia.org/wiki/Executable
3. Object file: https://en.wikipedia.org/wiki/Object_file
4. Library: https://en.wikipedia.org/wiki/Library_(computing)
5. Code segment: https://en.wikipedia.org/wiki/Code_segment
6. Data segment: https://en.wikipedia.org/wiki/Data_segment
7. .bss: https://en.wikipedia.org/wiki/.bss
8. Stack: https://en.wikipedia.org/wiki/Stack_(abstract_data_type)
9. Loader: https://en.wikipedia.org/wiki/Loader_(computing)

10. Dynamic loading: https://en.wikipedia.org/wiki/Dynamic_loading
11. Relocation: https://en.wikipedia.org/wiki/Relocation_(computing)
12. Position-independent code: https://en.wikipedia.org/wiki/Position-independent_code
13. Copy-on-write: https://en.wikipedia.org/wiki/Copy-on-write
14. Memory-mapped file: https://en.wikipedia.org/wiki/Memory-mapped_file
15. mmap: https://en.wikipedia.org/wiki/Mmap

See also:

- http://duartes.org/gustavo/blog/post/how-the-kernel-manages-your-memory/

阶段四：内核线程

- 学习Kthread
    - include/linux/kthread.h
    - kernel/kthread.c
    - 实例分析：kernel/hung_task.c
- 有趣的trick：
    - 设置程序argv[0]为0倒是可以骗过ps命令(ps显示的该进程命令为空)，但是骗不过top和/proc
    - 设置task_struct中的comm为空，top命令显示的进程名也空，但也能显示这个进程

其它建议内容：

- Process isolation: https://en.wikipedia.org/wiki/Process_isolation
    - Process isolation is a set of different hardware and software technologies[1] designed to protect each process from other processes on the operating system. It does so by preventing process A from writing to process B.
- Coscheduling: https://en.wikipedia.org/wiki/Coscheduling
    - Coscheduling is the principle for concurrent systems of scheduling related processes to run on different processors at the same time (in parallel). There are various specific implementations to realize this.
- Memory bound function: https://en.wikipedia.org/wiki/Memory_bound_function
    - Memory bound refers to a situation in which the time to complete a given computational problem is decided primarily by the amount of memory required to hold data. In other words, the limiting factor of solving a given problem is the memory access speed.
- Work stealing: https://en.wikipedia.org/wiki/Work_stealing
    - In parallel computing, work stealing is a scheduling strategy for multithreaded computer programs.
    - It solves the problem of executing a dynamically multithreaded computation, one that can "spawn" new threads of execution, on a statically multithreaded computer, with a fixed number of processor cores.
    - It does so efficiently both in terms of execution time and memory usage.
- Background process: https://en.wikipedia.org/wiki/Background_process
    - A background process is a computer process that runs "behind the scenes" (i.e. in the background) and without user intervention. Typical tasks for these processes include logging, system monitoring, scheduling, and user notification.
    - Daemon: https://en.wikipedia.org/wiki/Daemon_(computing)
        - In multitasking computer operating systems, a daemon (/ diːm n/ or / de m n/) is a computer program that runs as a background process, rather than being under the direct control of an interactive user.
        - Traditionally daemon names end with the letter d. For example, syslogd is the daemon that implements the system logging facility and sshd is a daemon that services incoming SSH connections.
- Ptrace: https://en.wikipedia.org/wiki/Ptrace
    - ptrace is a system call found in several Unix and Unix-like operating systems. By using ptrace (the name is an abbreviation of "process trace") one process can control another, enabling the controller to inspect and manipulate the internal state of its target.
    - ptrace is used by debuggers and other code-analysis tools, mostly as aids to software development.
- Thread-local storage: https://en.wikipedia.org/wiki/Thread-local_storage
    - Thread-local storage (TLS) is a computer programming method that uses static or global memory local to a thread.
    - TLS is used in some places where ordinary, single-threaded programs would use global variables but where this would be inappropriate in multithreaded cases.