# Topic 1: Basic Concept

## Overview

### Guidance

* Chapter: Operating-System-Concepts.pdf 1, 2.2, 2.3; Unix-Network-Programming-Vol.2 Part 1
* Spend time: 2 days
* Learn suggestion: Take a general understanding of the operating system
* Key points: POSIX, System Call, Signal, Timer, File Descriptor
  + What's the OS? Why do we need OS?
  + The history of Unix, What's the relationship between Linux and Unix?
  + What's the fundamental function of OS?
  + Try to get help from the Linux help manual
  + Write some programs for each type of system call using the function in page 68.

### Practice

None

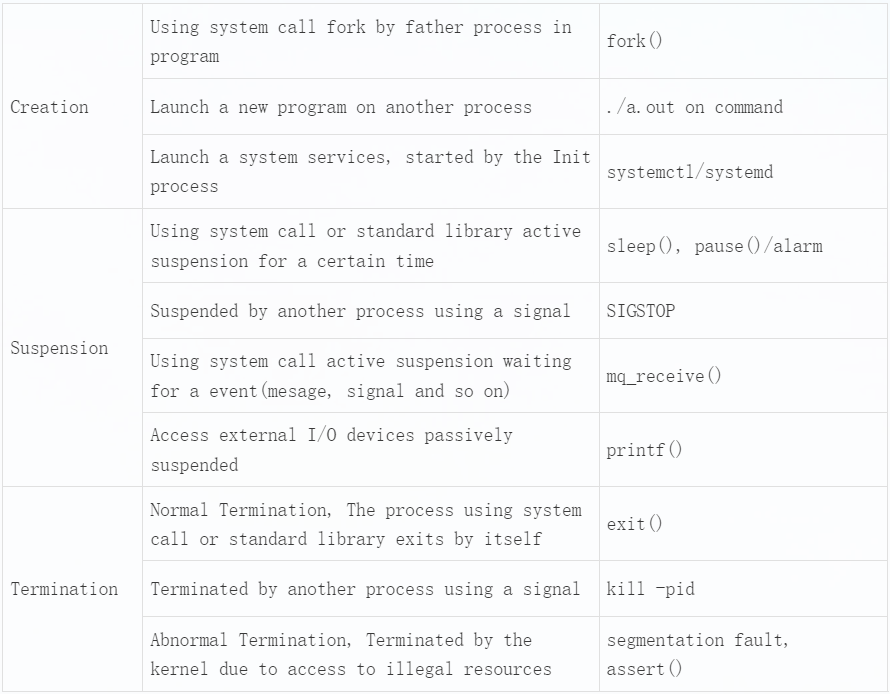
## Process

### Guidance

* Chapter: Operating-Systems-Internals-and-Design-Principles.pdf 3
* Spend time: 1 day
* Learn suggestion: To know about PCB, Process state, Process creation/termination
* Key points:
  + Using command ps/top to check the information of process, Try to explore what options can be used
  + How to capture the different process state?
  + The memory layout of Process address space

### Practice

1. How many approaches can a process be created, suspended, or terminated? Illustration



1. What's the meaning of the zombie and orphan process? Create a zombie and orphan process separately and check it by command ps

Orphan process: When a parent process exits while one or more child processes are still running, they become orphan processes. Orphan processes are adopted by the init process (process id 1), and the init process collects status on them.

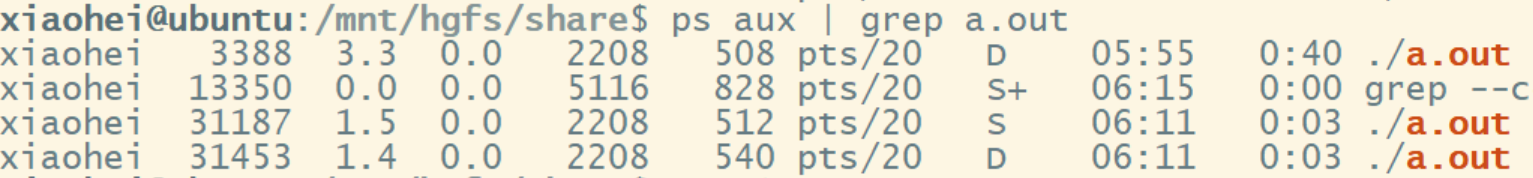
Zombie process: A process creates a child process using fork. If the child process exits and the parent process does not call wait or waitpid to get the state information of the child process, the process descriptor of the child process is still saved in the system. This process is called a dead process.



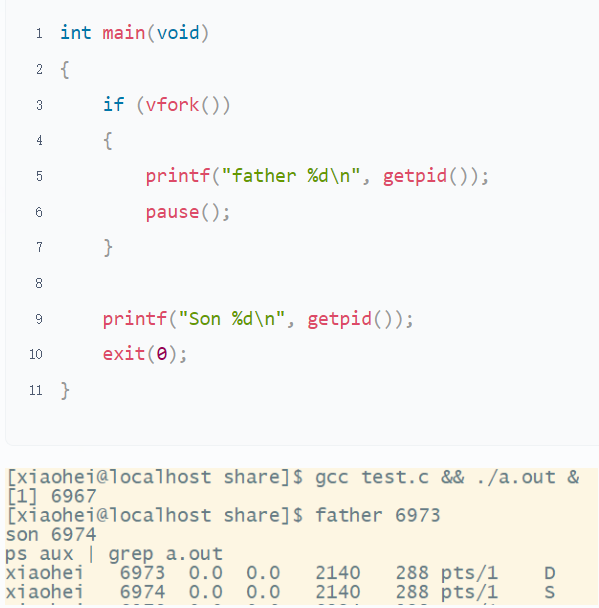


1. Try to get the state TASK\_UNINTERRUPTIBLE(D) of process in linux
   1. Prolongs the waiting time of a process in the task queue





* 1. vfork(Not uninterruptible)



* 1. Using kmod to lock kernel processes.(Leave it for you to explore)

## Thread

### Guidance

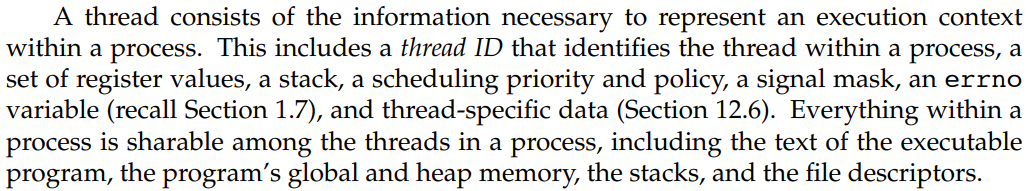
* Chapter: Operating-System-Concepts.pdf 4
* Spend time: 1 day
* Learn suggestion: Learn to compare with process and coroutine
* Key points:
  + Difference between process and thread
  + Difference between ULT and KLT
  + How many ways to implement the ULT?

### Practice

1. For multiple threads in the same process, which resources are shared and which are unique?

* Resources exclusive by thread

Advanced-Programming-in-the-UNIX-Environment P284



note: The global variable can become thread exclusive after declared by \_\_thread.

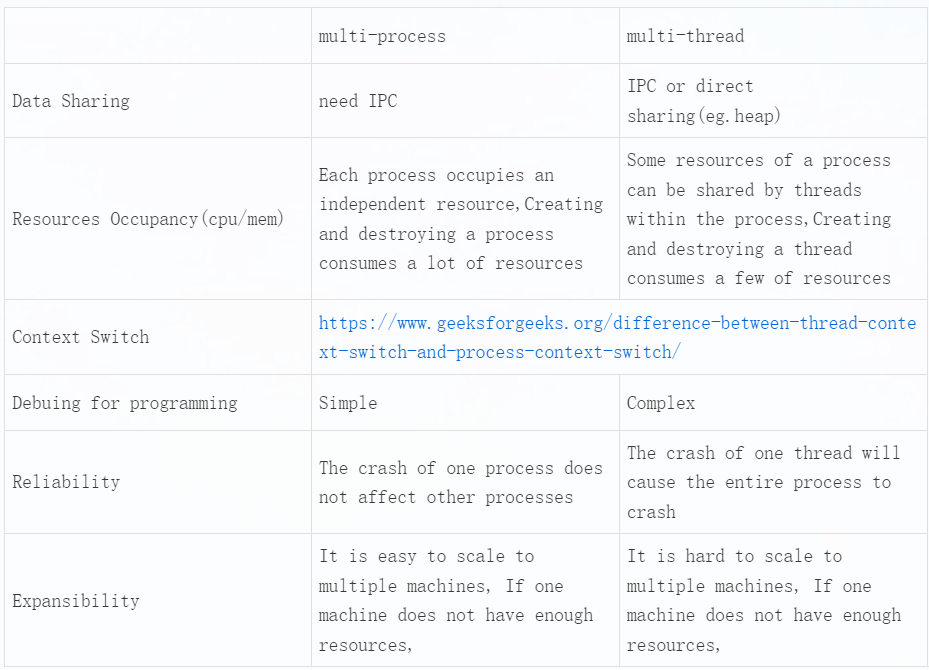
* Resources shared by threads

Global variable table(.text, .data, and so on), heap, dynamic-link library, files, Current working directory, as well as user ID and group ID

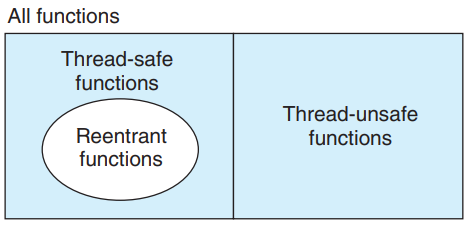
1. What are the advantages and disadvantages of multi-process(only one thread in each process) and multi-thread programming?

Process: Minimum unit of resource allocation

Thread: Minimum unit of cpu scheduling



1. What are reentrant and non-reentrant functions? What are thread-safe and thread-unsafe functions? Illustration



Reentrant functions and Thread-safe functions are different concepts. Refer to p1059 on "Computer-Systems-A-Programmers-Perspective.pdf"

Reentrant functions: Refer to chapter 10.6/12.5 on "Advanced-Programming-in-the-UNIX-Environment.pdf"

Thread-safe functions: When a function is called by different threads at the same time, each thread will get the correct result. This can only occur in multi-threaded scenarios.

Reentrant functions: If a function is interrupted at any time during execution, it can still be reused, or the correct result can be obtained after the interrupt is returned. This can occur in multi-threaded or single-threaded scenarios.

Determine non-reentrant functions

* Functions that directly use or return global/static variables, including the errno variable
* malloc/free is called
* Standard I/O function is called， most of which use global data structures in a non-reentrant manner
* Other non-reentrant functions

A non-reentrant function can be locked to become thread-safe, but it's not reentrant. Reentrant functions are higher-level thread-safe functions, a concept similar to asynchronous signal/interrupt safety functions.

The function is a reentrant if its data is in its own stack space.

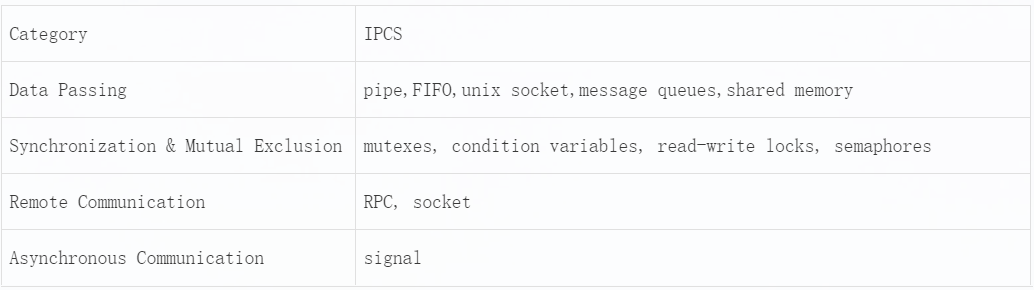
## IPC

### Guidance

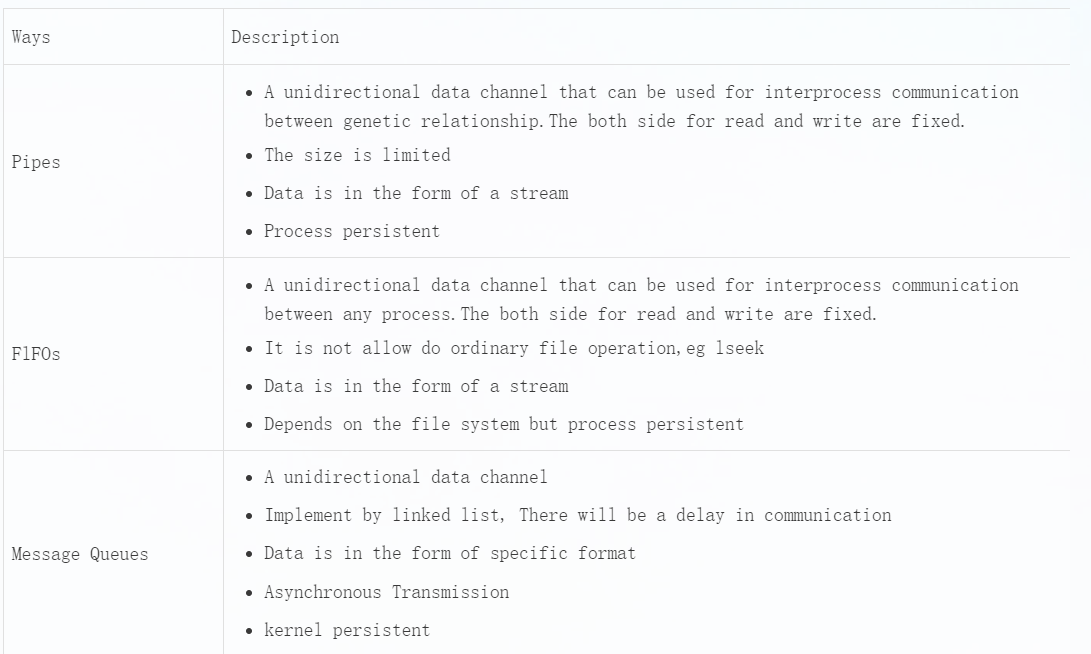
* Chapter: Unix-Network-Programming-Vol.2.pdf Part 2-5; Advanced-Programming-in-the-UNIX-Environment chapter 17
* Spend time: 2 days
* Learn suggestion: Compare and understand the characteristics of communication modes between different processes
* Key points:
  + How many ways to communicate between different processes?
  + The life cycle of the IPC object
  + What is different between Posix and System V IPC?
  + Can these ways be used for inter-thread communication?
  + What are the characteristics of each IPC?

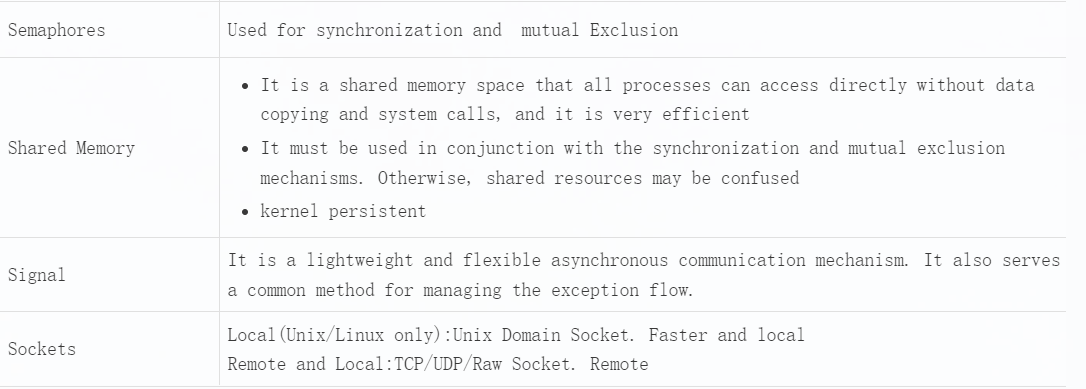
### Practice

1. According to your understanding, try to divide all IPCs into different categories, and the reason



1. To summarise, what are the characteristics of each IPC?





1. Using fork/exec to simulate executive "ping 127.0.0.1 -c 3" and get the result

Refer to "fork-exec simulate executive ping.c"

1. Write a program to implement the synchronisation and mutual exclusion by semaphore or lock

Refer to "Synchronisation and Mutual Exclusion.c"

1. Follow the client/server model: client sends a filename to the server, then the server responds its content to the client(Unix Domain Socket)

Requirements:

* Return the state of server to client when the specified file can't be opened correctly
* Follow the TLV format to organise data
* Don't mix the file operation and socket operation together. Separate them

Refer to "UDP based domain socket"

## Signal

### Guidance

* Chapter: Advanced-Programming-in-the-UNIX-Environment.pdf 10.1-10.10
* Spend time: 1.5 days
* Learn suggestion: Have a general understanding of the signal and where it's being used. Don't need to learn the advanced usage of it for now.
* Key points:
  + How many signals can be used? What's the meaning of each signal?
  + How to check all signals in Linux? Where is the signal 32/33?
  + What's the difference between signal and interrupt?
  + What's a real-time signal and a non-real-time signal?
  + How do you control the creation of core dump files when a program crashes?

### Practice

1. Modify the default behavior of ctrl + c when ctrl + c is pressed, the output is Hello World! ; To achieve the same effect by the command kill



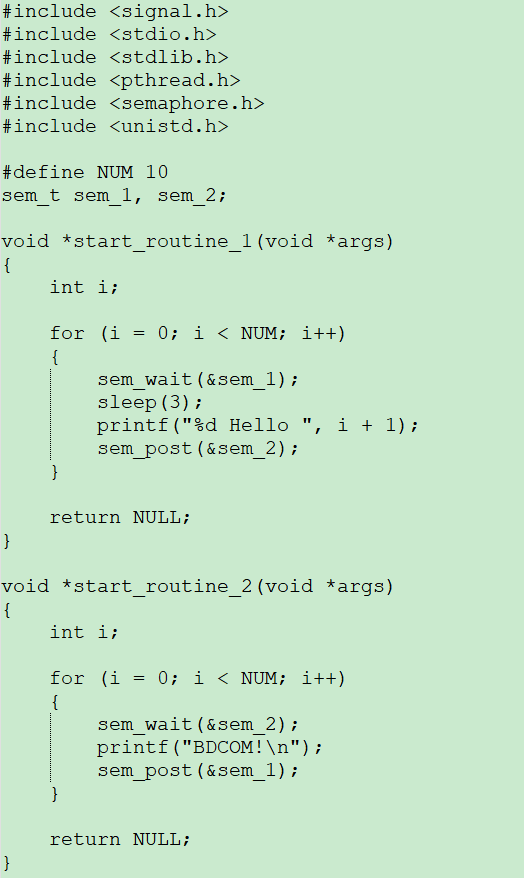
1. Write a program to simulate the try/catch/finally in C. It's able to handle both file inexistence and segment fault

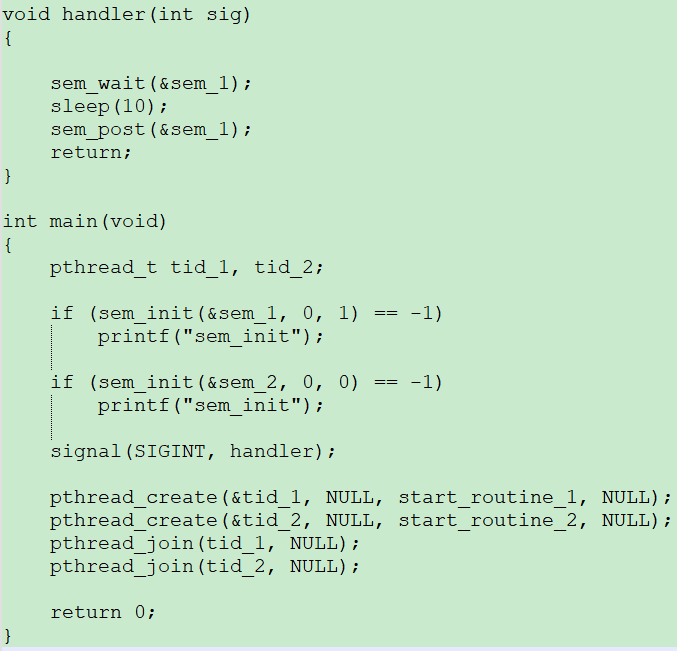
Refer to "Implement to simulate the try-catch-finally.c"

1. Implement the function sleep() and a periodic timer using alarm() and pause() simulations. And wake up the process from the sleep state in advance in the timer periodically

Refer to "Implement to simulate sleep and alarm by signal.c"

1. What issues does signal may bring to the program? Illustration
   1. Disturbs the program execution sequence
   2. Disrupts program logic, causing resources to wait for each other





* 1. A non-reentrant function is called in both a normal function and a signal handler at the same time, which may lead to a data mess