Second laboratory task

General information

The information system consists of three types of nodes (those that serve requests (P), those that organize request queues (Q) and transfer requests for service (P), and mixed nodes (QP), i.e., those that have both a request queue and request service). TCP/IP (socket) is used for communication between components.

The application queue (Q) can have the following properties (configurable parameters):

- Single priority, organized according to FiFO or LIFO principles;
- Multiple priorities, i.e., applications with higher priority (mark) are given priority for service;
- Removes the application from the queue only after service (ensures reliability);
- Removes the request from the queue as soon as it is transferred for processing;
- If requests of the same type are processed by several processing devices, the following request transfer strategies may be applied:
    - First available from the list of devices;
    - Cyclically, i.e., to the next available servicing device (pseudo-balancing between servicing devices) when there are several available servicing devices. In the case of all available devices, $j = i \bmod m + 1$, where m is the number of serving devices, *i* is the index of the last serving device that "received" the request, and *j* is the index of the device that will serve the request.
- Require a reduction in the length of the queue for additional service facilities according to certain criteria:
    - When the length of the queue exceeds the set limit;
    - When the maximum application retention time exceeds the set time;
    - When the average application retention time in the queue exceeds the set time.

The service facility serves the request within a certain time interval:

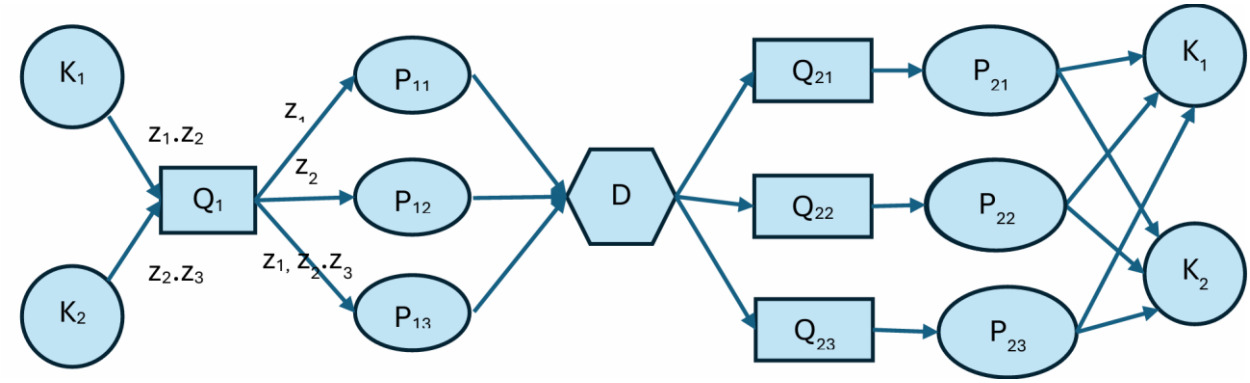
- Fixed
- Randomly distributed according to a certain law (uniformly distributed in the interval $[a, b]$, exponential $\delta = \frac{1}{\lambda}\chi, \chi$ – uniformly distributed random variable in the interval [0,1], normal $\delta = \sqrt{-2\ ln\ \chi_1}\ cos\ 2\pi\chi_2, \chi_1, \chi_2$ - uniformly distributed random variables in the interval [0,1]

All system nodes (processes) operate asynchronously and may be destroyed if the servicing device is not used for too long or breaks down.

There are a number of clients (K) that can send requests with tags (z) to the system for servicing and receive a result, even in the event of node failure.

Assignment task

An imitation model of the system must be implemented for the diagram below with specific parameters.



z_3 has priority over z_2, z_2 has priority over z_1. *D* – high-performance query distribution device. The client must receive responses from all three P_{2x} service devices within a specified time and counts the number of successful and failed requests. Returned processed requests are successful if the number of received processed requests is the same as sent out number (e.g. z_1, z_2 has been sent out and returned back after processing) and vice versa (e.g. due to fail of some P_{2x} devices, not all request had been returned back, then we have *failed requests* case). P_{2x} – may get down. Q1 distributes requests cyclically. $K_{1,2}$ on left and right are the same units.

Dynamic change of configuration: in case of increased average time while the specific request (z_1 *and/or* z_2 *and/or* z_3) is in a queue, the number of associated with that request processing unit should be increased. E.g. requests of z_3 type average awaiting time is greater than usual, then number of P_{13} units should be increased.

Grading

- Correct formal specification – 2 points,
- implementation of the system simulation model – 6 points,
- introduction of dynamically changing elements into the system – 2 points.

Submission deadline: End of November.