Вероятностное моделирование поведения BOINC-инфраструктуры в типовых ситуациях

Докладчик: Храпов Николай Павлович

Реализация BREB в BOINC

- Фактическое время между запросами равно максимальному времени ожидания, умноженному на случайное число от 0 до 1. Максимальное время ожидания после первого необслуженного запроса равно 10 минутам.
- Максимальное время между запросами удваивается при очередном необслуженном запросе.
- Интервал максимального времени между запросами будет экспоненциально возрастать до значения в 24 ч. После достижения такого значения возрастание прекратится, и все последующие максимальные интервалы ожидания составят 24 ч.

Поток запросов после деактивации сервера

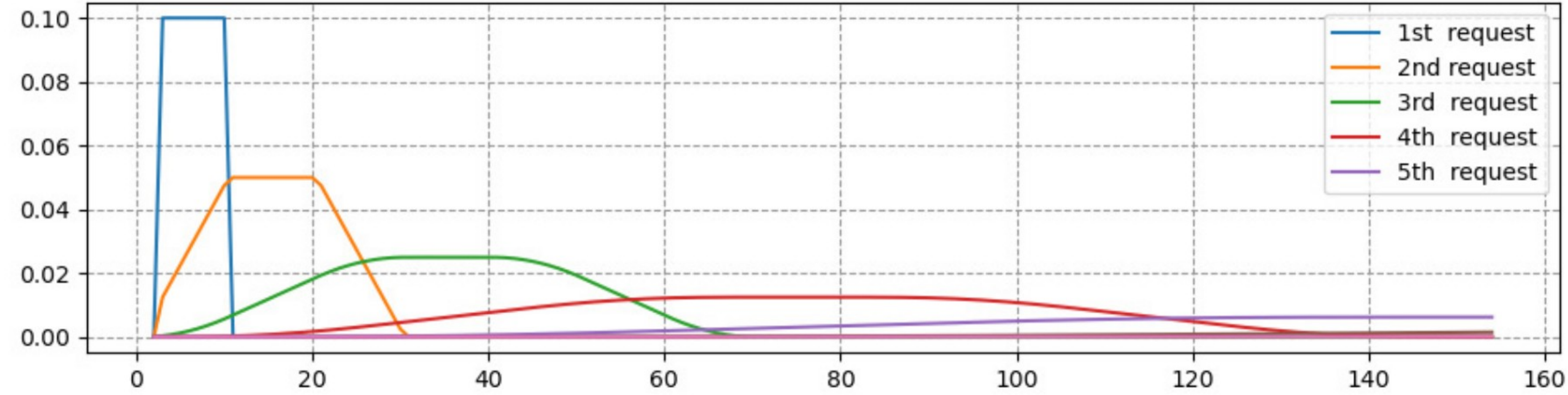
Рассмотрим следующую ситуацию:

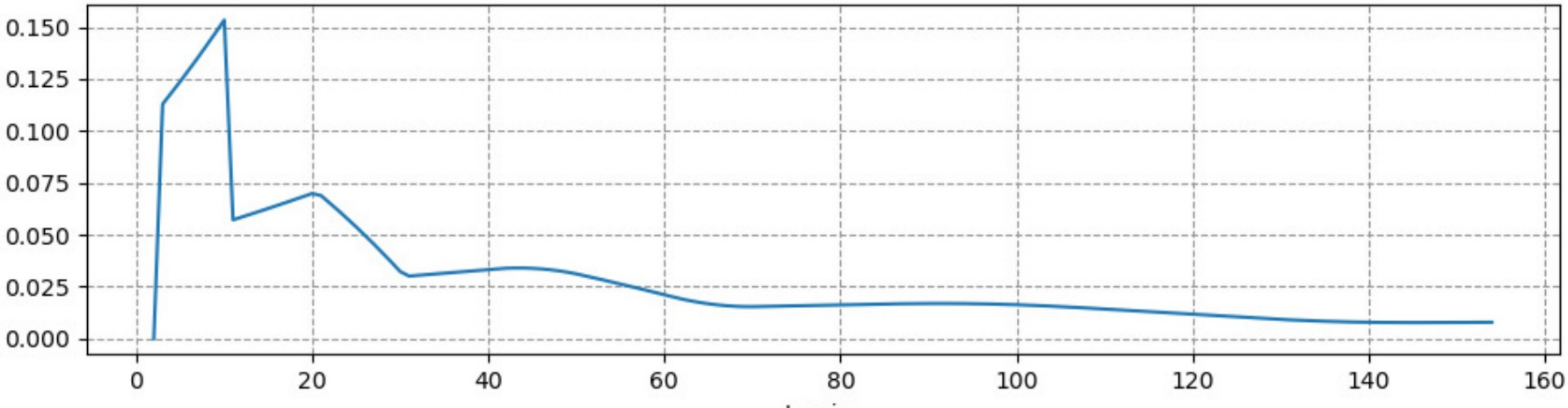
- Инфраструктура состоит из N стабильно работающих вычислительных узлов и выполняет задания продолжительностью T_{task}.
- Поток заданий в единицу времени при стабильно работе:

$$I = \frac{N}{t_{task}}$$

- Необходимо посчитать зависимость плотности потока запросов от времени после остановки проекта.

Плотность вероятности выполнения узлом 1, 2, 3, 4, 5-го запроса через t минут после обнаружения деактивации и суммарная плотность веротностей.





Поток запросов после деактивации сервера

Время обнаружения деактивации будет различным для разных узлов.

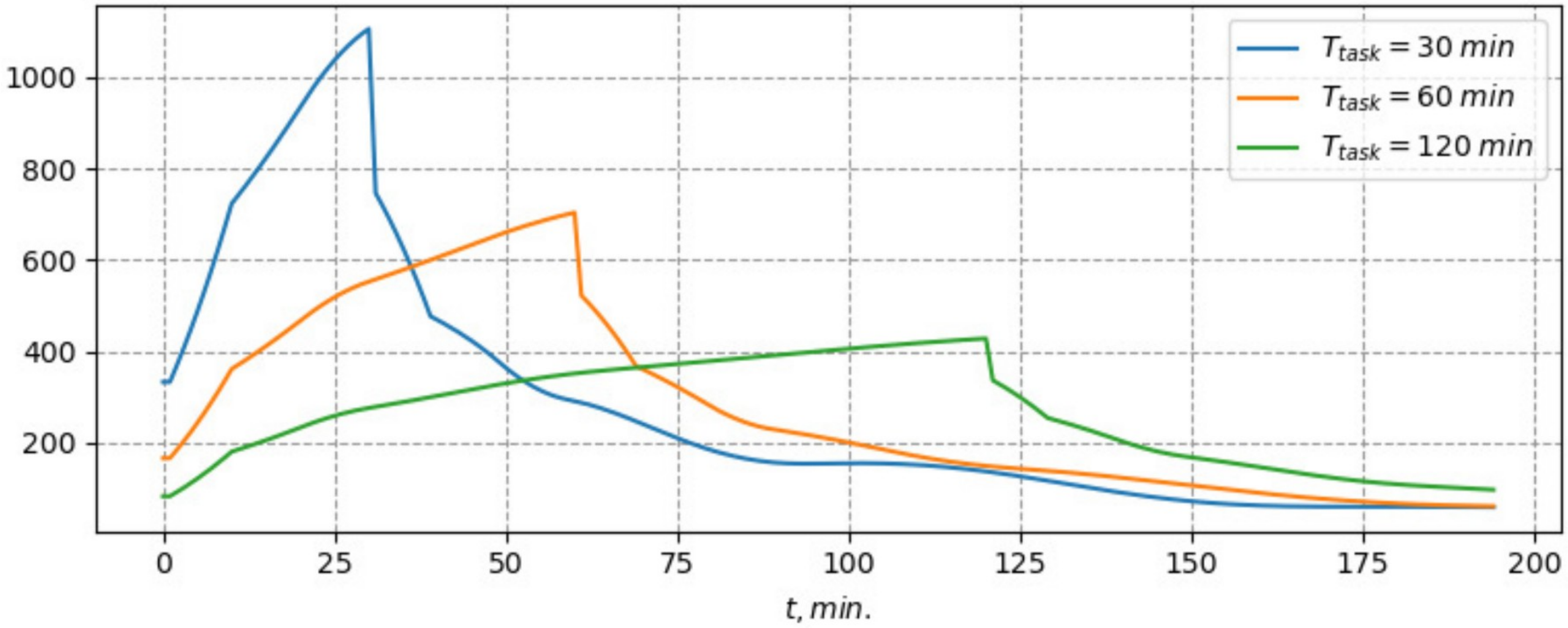
- dN узлов сделали запрос в период $[\tau_{initial};\tau_{initial} + dt]$.
- Вклад этих dN узлов в поток запрос будет составлять

$$dI = P_{total}(t - \tau_{initial})dN$$

- Тогда поток запросов через t минут после деактивации составит

$$I(t) = \frac{N}{T_{limit}} \int_0^{T_{limit}} P_{total}(t - \tau_{initial}) d\tau_{initial} \quad , \qquad T_{limit} = \begin{cases} t, & if \;\; t < T_{task} \\ T_{task}, & if \;\; t \geq T_{task} \end{cases}$$

Динамика потока запросов после деактивации проекта

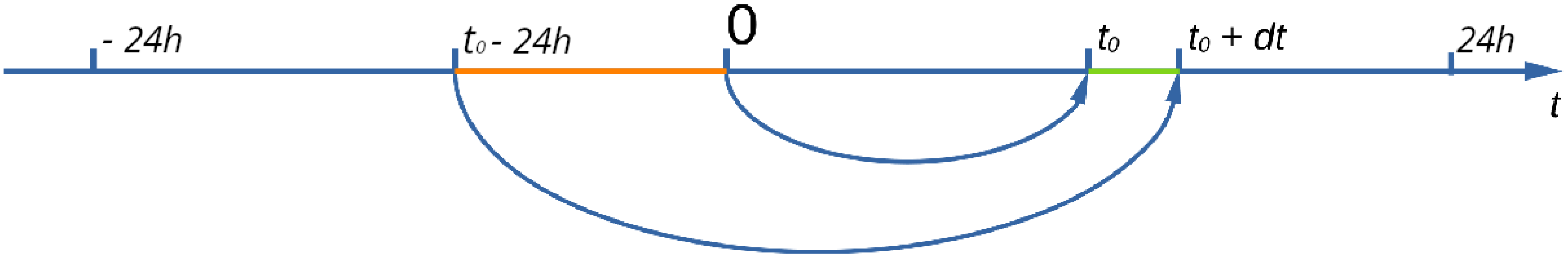


Поток первичных запросов после активации сервера

Деактивацию сервера будем считать продолжительной, если её продолжительность более 2-х суток. За это время поток запросов пример стационарный характер.

Вероятность выполнения запроса в период [t; t + dt] пропорциональна длине интервала выполнения последнего необслуженного запроса

$$P\{t \in (t_0; t_0 + dt)\} \sim T_{max} - t_0$$



Плотность вероятности выполнения запроса убывает линейно после активации сервера

$$p(t) = p_{max} - \frac{p_{max}}{T_{max}} \cdot t = \frac{2}{T_{max}} - \frac{2}{T_{max}^2} \cdot t$$

$$p_{max} = \frac{2}{T_{max}} = \frac{2}{24 \cdot 60}$$

Поток первичных запросов пропорционален плотности вероятности.
Доля узлов, активировавшихся к моменту t:

$$A(t) = \int\limits_0^t p(\xi) d\xi = p_{max} \cdot t - \frac{p_{max}}{2 \cdot T_{max}} \cdot t^2 = \frac{2}{T_{max}} \cdot t - \frac{1}{T_{max}^2} \cdot t^2$$

Инерциальные потери вычислительной инфраструктуры составят

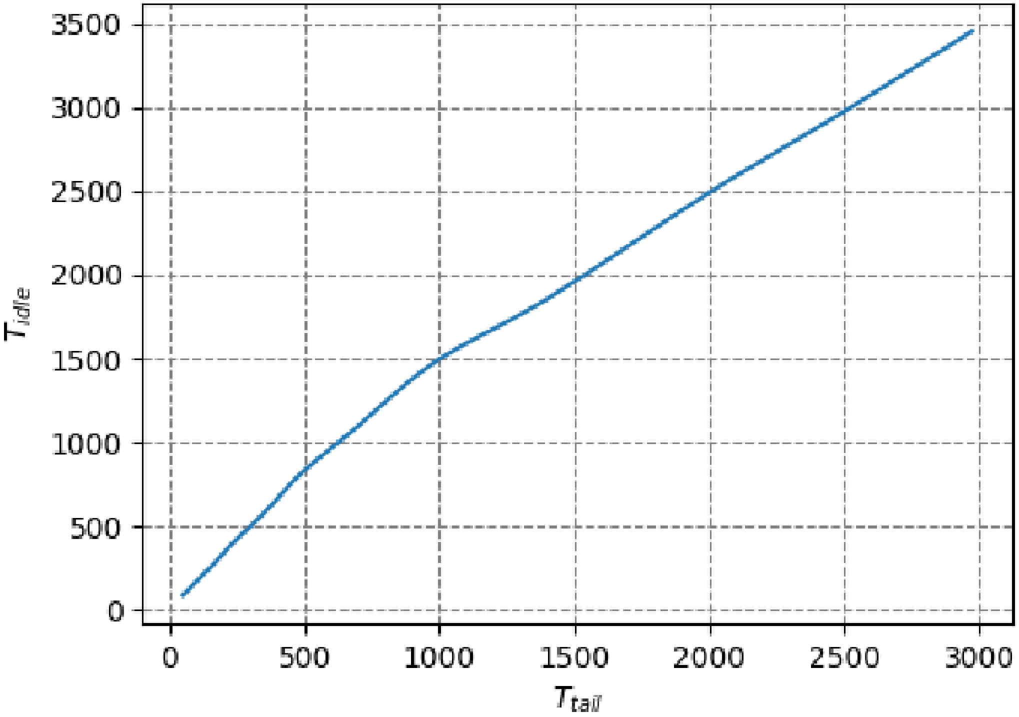
$$L = \int\limits_0^{T_{max}} 1 - A(t) dt = \int\limits_0^{T_{max}} 1 - \frac{2}{T_{max}} \cdot t + \frac{1}{T_{max}^2} \cdot t^2 dt = \frac{T_{max}}{3} = 8\ h.$$

Временная деактивация проекта

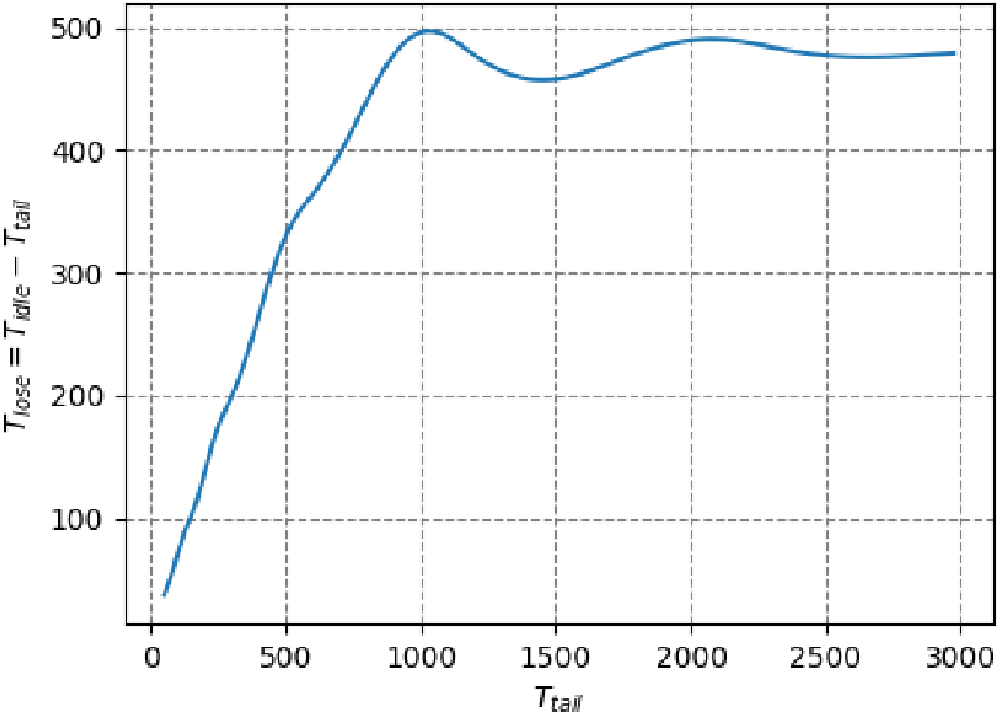
Рассмотрим следующую ситуацию:

- Инфраструктура состоит из N стабильно работающих вычислительных узлов и выполняет задания продолжительностью T_{task}.
- Поток заданий в единицу времени при стабильно работе: $I = \frac{N}{t_{task}}$.
- В некоторый момент времени происходит деактивация проекта продолжительностью $T_{deactive}$.
- Необходимо оценить суммарные потери процессорного времени инфраструктуры, вызванными временной деактивацией.

Зависимость математического ожидания времени бездействия (а) и потерянного времени (б) от времени обнаружения деактивации T_{tail}



а)



б)

Количество узлов, осуществивших запрос в период *dt*:

$$dN = N \cdot \frac{dt}{T_{task}}$$

Усреднённые потери составят $dL = M(t)\, dN$.

Суммарные потери составят:

$$L = \int_{T_{deactiv}-T_{task}}^{T_{deactiv}} \frac{N}{T_{task}} \cdot M(t)dt = \frac{N}{T_{task}} \cdot \int_{T_{deactiv}-T_{task}}^{T_{deactiv}} M(t)dt$$

Рассмотрим пример. Инфраструктура, состоит из 10 тыс. стабильно работающих узлов. Время выполнения задания при стабильной работе составляет 1 час.

Время деактивации составляет 4 часа.

Тогда время обнаружения дективации для узлов равномерно распределено в интервале от

3-х до 4х часов.

На этом участке мат. ожидание имеет квазилинейный вид и может быть представлено следующим образом:

$$M(T_{tail}) = 1.65 \cdot T_{tail} + 5.23$$

Тогда суммарные потери составят:

$$L = \frac{N}{T_{task}} \cdot \int\limits_{T_{deactiv}-T_{task}}^{T_{deactiv}} M(t)dt = \frac{10000}{60} \int\limits_{180}^{240} (1.65t + 5.23)dt =$$

$$= 3517300 min. = 58622 h.$$

??????????????

Плотность вероятности выполнения запроса убывает линейно после активации сервера

$$p(t) = p_{max} - \frac{p_{max}}{T_{max}} \cdot t = \frac{2}{T_{max}} - \frac{2}{T_{max}^2} \cdot t$$

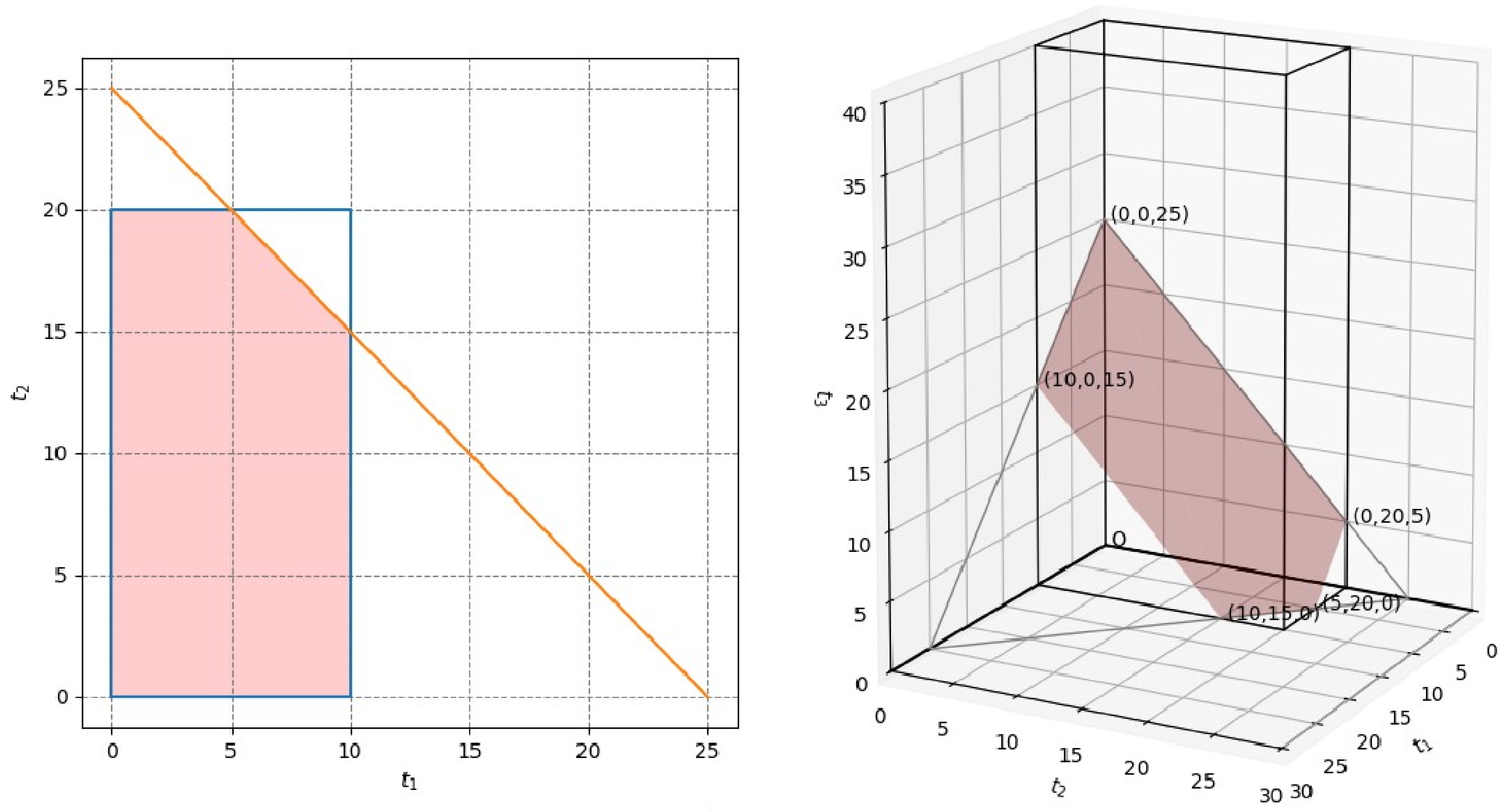
$$p_{max} = \frac{2}{T_{max}} = \frac{2}{24 \cdot 60}$$

Поток первичных запросов пропорционален плотности вероятности.
Доля узлов, активировавшихся к моменту t:

$$A(t) = \int_0^t p(\xi) d\xi = p_{max} \cdot t - \frac{p_{max}}{2 \cdot T_{max}} \cdot t^2 = \frac{2}{T_{max}} \cdot t - \frac{1}{T_{max}^2} \cdot t^2$$

Инерциальные потери вычислительной инфраструктуры составят

$$L = \int_0^{T_{max}} 1 - A(t) dt = \int_0^{T_{max}} 1 - \frac{2}{T_{max}} \cdot t + \frac{1}{T_{max}^2} \cdot t^2 dt = \frac{T_{max}}{3} = 8\ h.$$



t_2
t_1
(0,0,25)
(10,0,15)
(0,20,5)
(10,15,0)
(5,20,0)
t_3

Построение функции распределения

Функция распределения равна вероятности, что при фиксированном T_{tail} время бездействия системы будет меньше t:

$$F(T_{tail}, t) = P\{T_{tail}, T_{idle} < t\}$$

$$F(T_{tail}, t) = \sum_{i=1}^{\infty} P\{A_i' \cap \{T_{idle} < t\}\} = \sum_{i=1}^{\infty} F_i(T_{tail}, t)$$

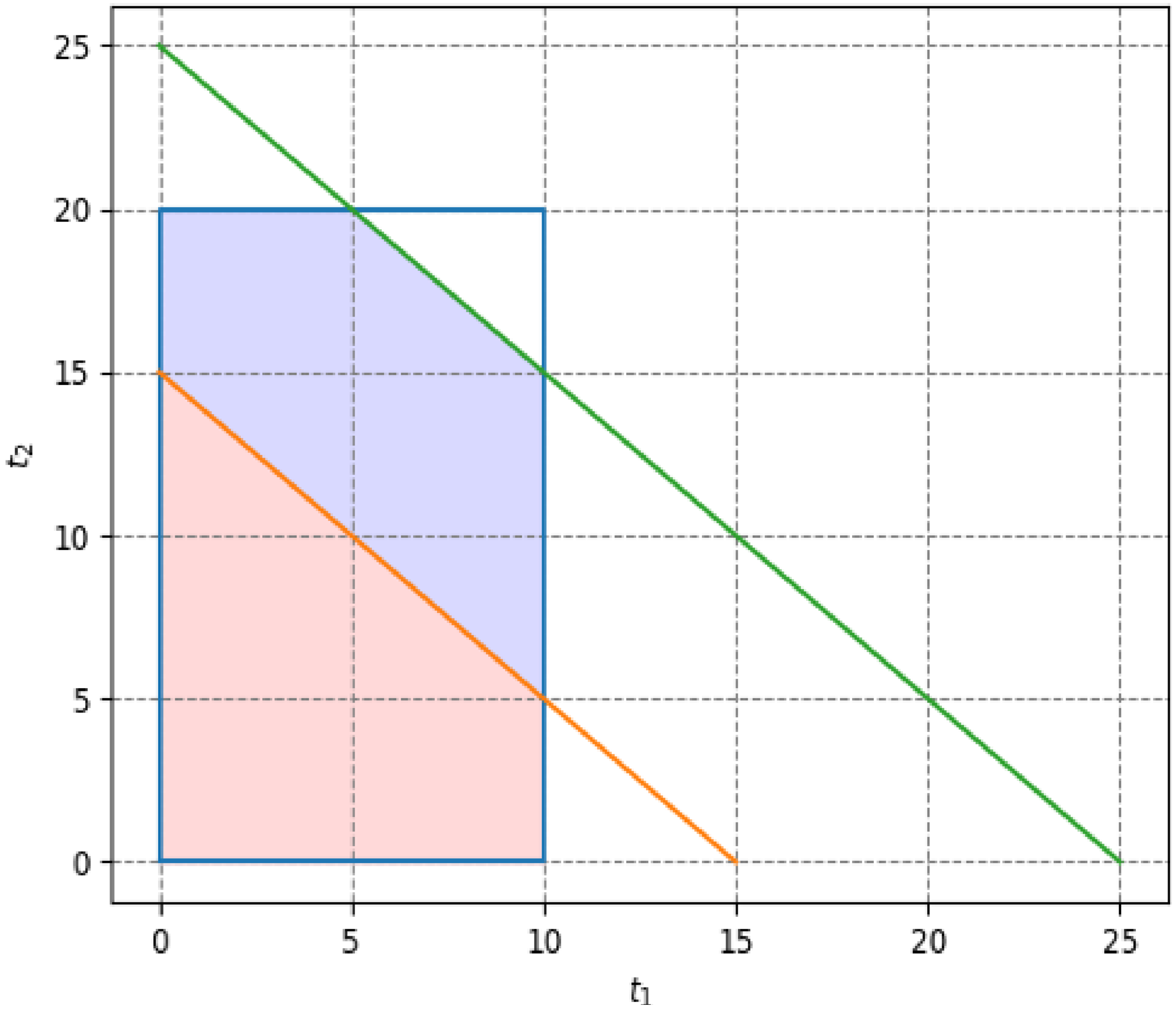
Вероятность, что при фиксированном T_{tail} и при фиксированном числе запросов время бездействия системы будет меньше t:

$$F_i(T_{tail}, t) = P\{A_i' \cap (T_{idle} < t)\}$$

Оценка вероятности события, что при

T_{tail} = 15 минут

будет выполнено ровно 2 запроса и суммарное время бездействия составит менее 25 минут.

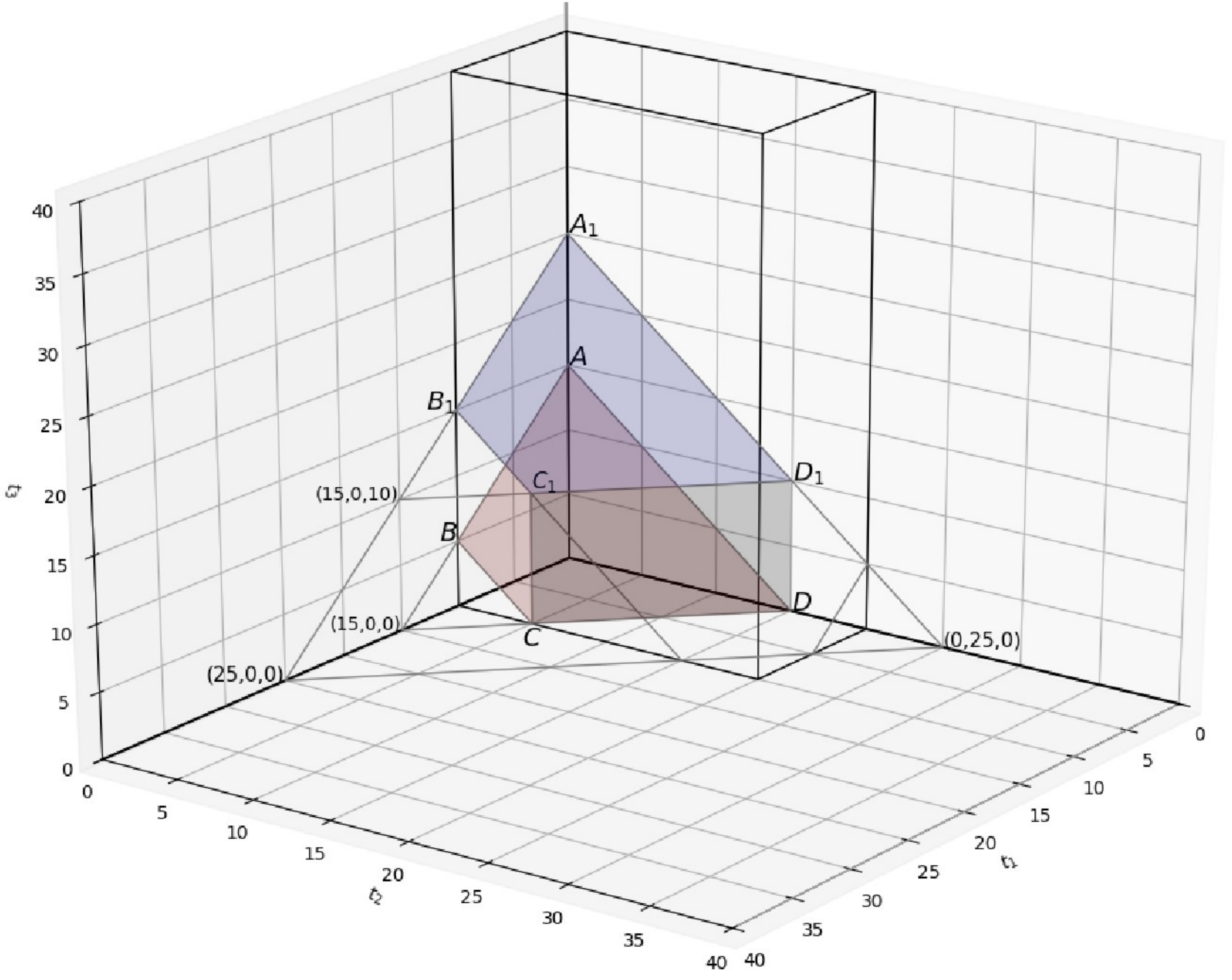


Оценка вероятности события, что при

T_{tail} = 15 минут

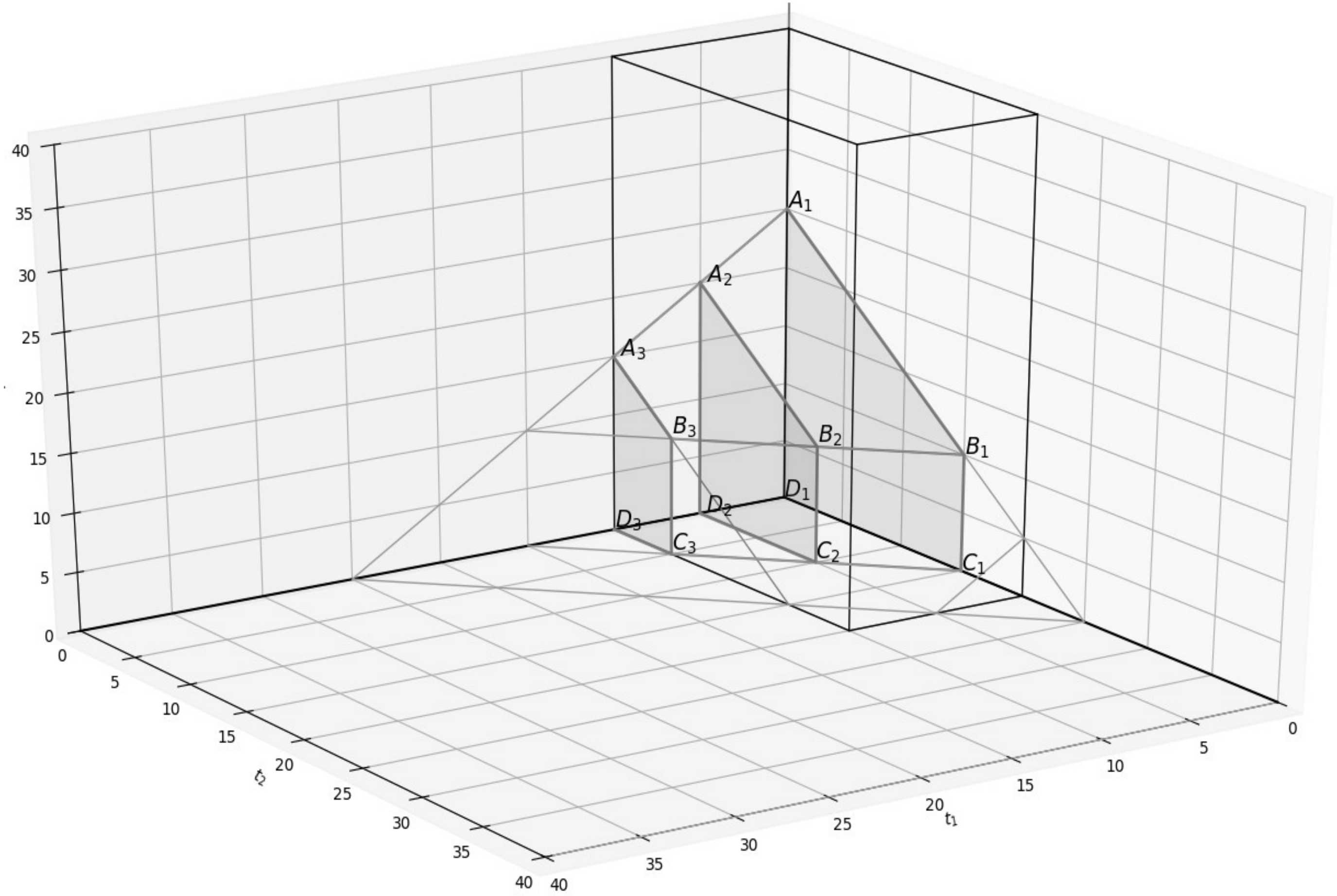
будет выполнено ровно 3 запроса и суммарное время бездействия составит менее 25 минут. Ограничения удовлетворяют условиям:

$$\begin{cases} \sum_{i=1}^{n-1} \tau_i < T_{tail} \\ T_{tail} < \sum_{i=1}^{n} \tau_i < t \\ 0 < \tau_i < 10 \cdot 2^i \end{cases}$$

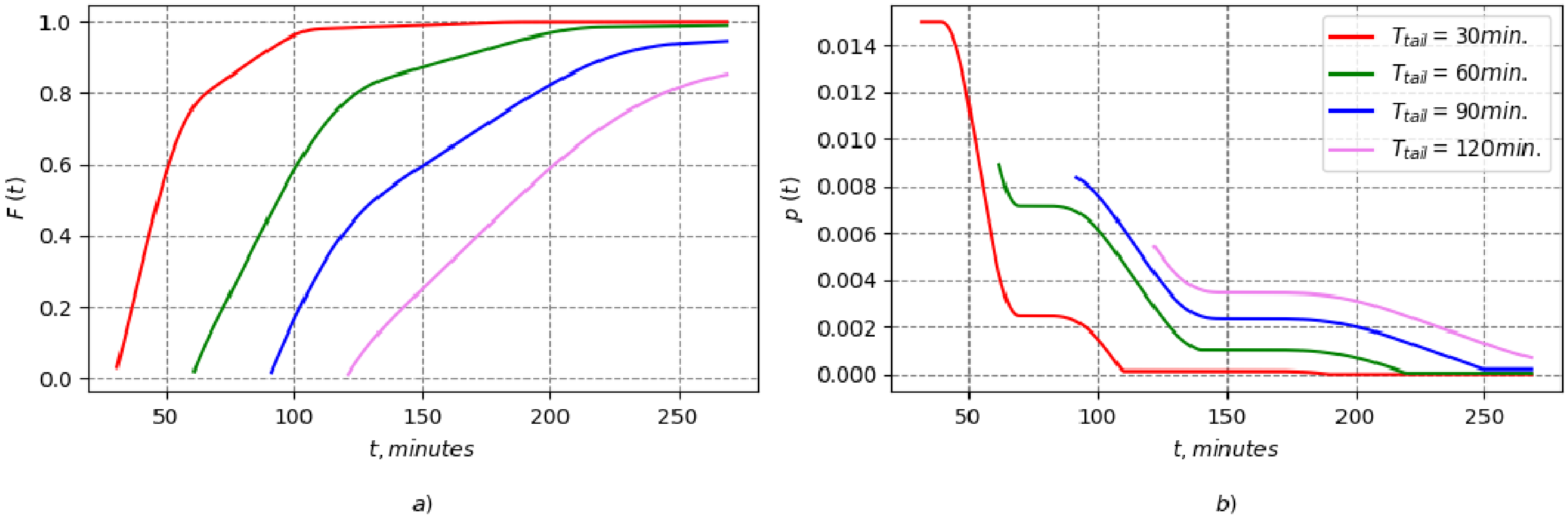


Пример применения рекурсивного способа вычисления объёма многомерной фигуры, удовлетворяющей ограничениям

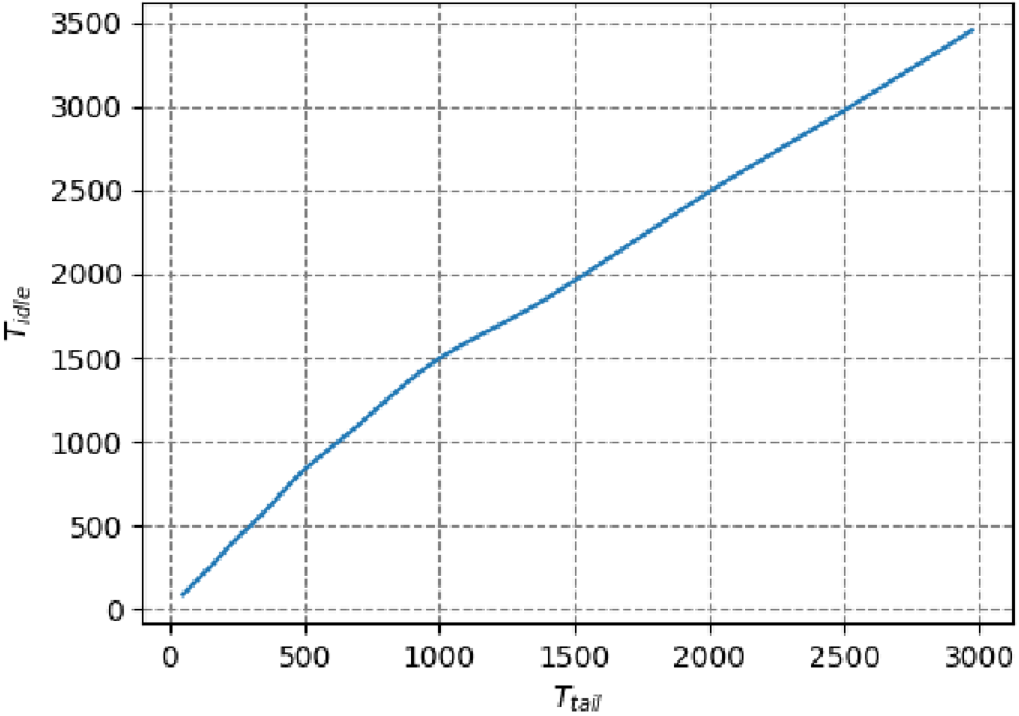
$$\begin{cases} \sum_{i=1}^{n-1} \tau_i < T_{tail} \\ \sum_{i=1}^{n} \tau_i < t \\ 0 < \tau_i < 10 \cdot 2^i \end{cases}$$



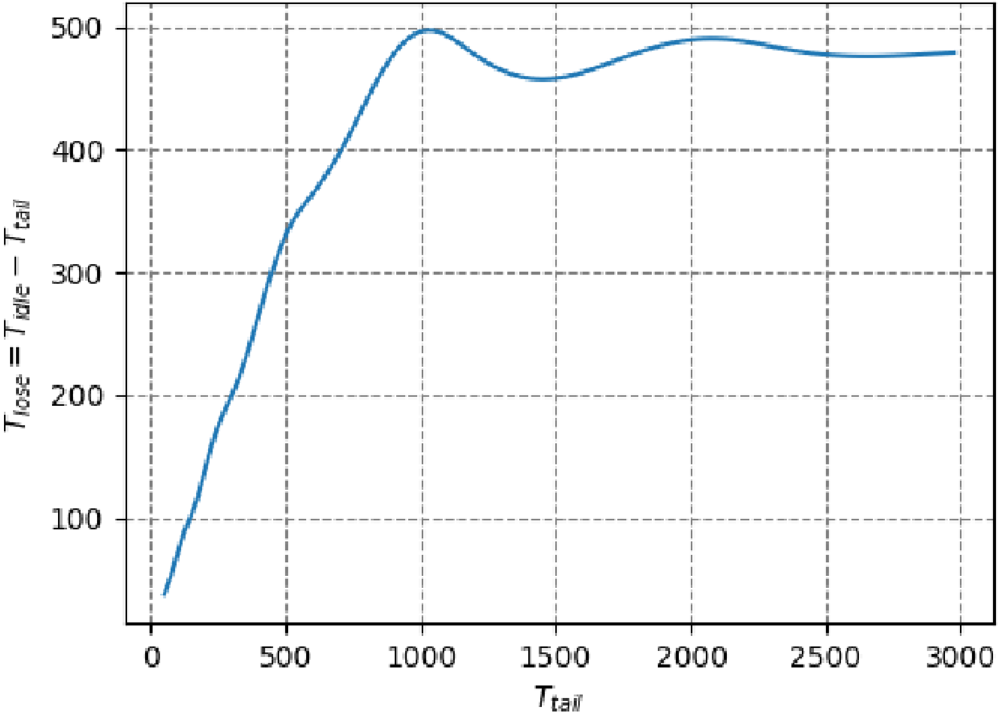
Функция распределения (а) и плотности вероятности(б)



Зависимость математического ожидания времени бездейсвия (а) и потерянного времени (б) от T_{tail}



а)



б)