

# Предмет, цель и задачи теории массового обслуживания

*Системой массового обслуживания (СМО)* называется любая система, предназначенная для обслуживания каких-либо заявок (требований), поступающих на нее в случайные моменты времени.

*Теория массового обслуживания (ТМО)* изучает случайные процессы, протекающие в системах массового обслуживания.

*Предметом теории массового обслуживания* является количественная сторона процессов, связанных с массовым обслуживанием. Для анализа этих процессов строится *математическая модель* обслуживающей системы, связывающая заданные условия работы системы с показателями эффективности, описывающими ее способность справляться с потоком требований. При исследовании системы массового обслуживания используются методы дискретной математики теории дифференциальных уравнений, теории вероятностей и математической статистики, теории случайных процессов.

*Основными задачами теории массового обслуживания являются:*

- построение математической модели системы массового обслуживания и расчет её основных характеристик;  
Математическая модель - это система математических соотношений, приближенно (в абстрактной форме) описывающих изучаемую систему. При построении математической модели отбрасываются несущественные с точки зрения исследователя факторы, влияющие на работы системы.
- установление зависимости эффективности работы системы массового обслуживания от её организации, выявление проблемных мест в ее структуре и организации работы;
- решение различных оптимизационных задач, связанных с функционированием системы массового обслуживания.

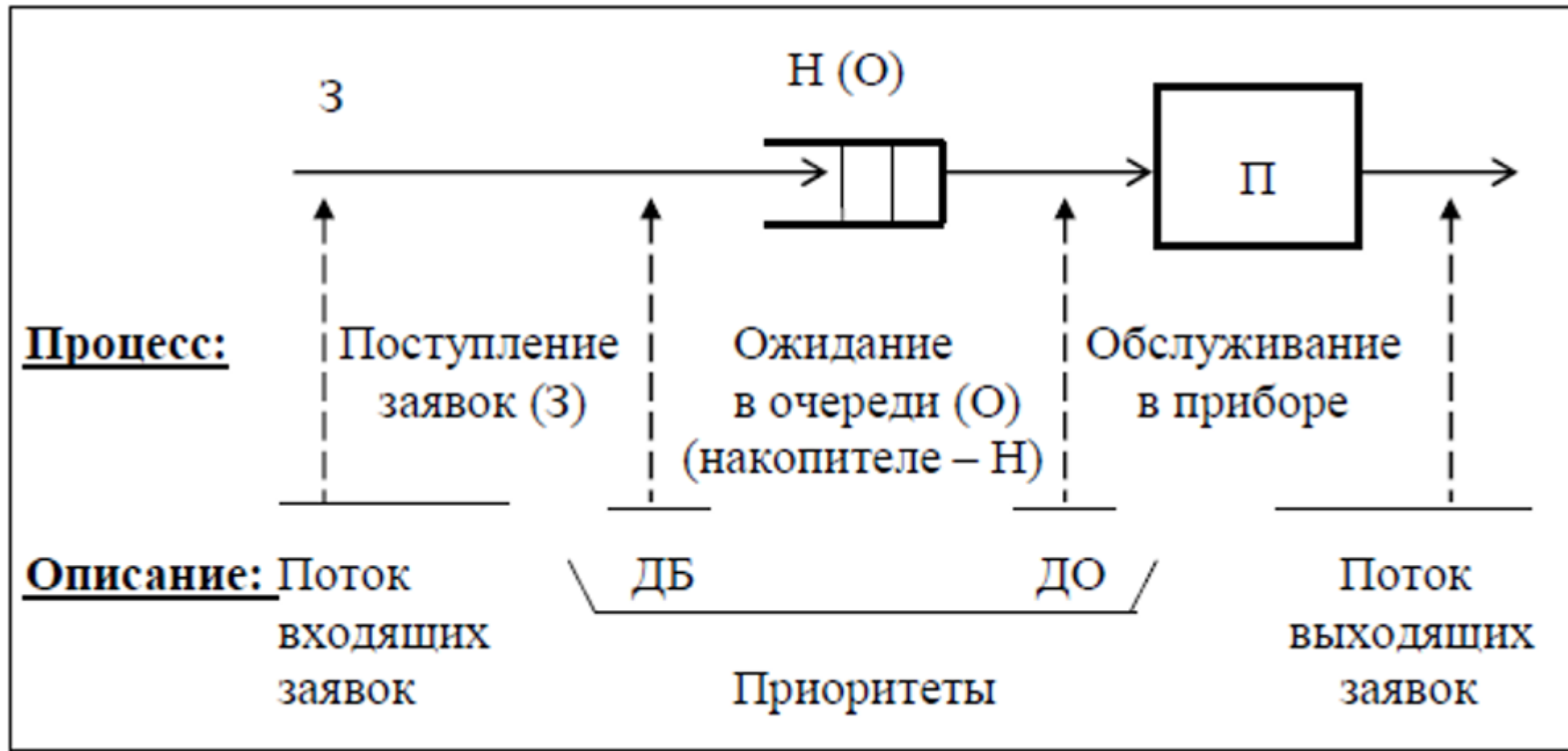
*Целью теории массового обслуживания является выработка рекомендаций по рациональному построению обслуживающих систем, рациональной организации работы систем для обеспечения высокой эффективности их функционирования.*

## Характеристика исторических этапов становления и развития теории массового обслуживания

Этап	Ученые, внесшие вклад в развитие	Краткая характеристика этапа
Возникновение теории вероятностей (середина XVII в.)	Б. Паскаль П. Ферма Х. Гюйгенс	Исследования в области азартных игр привели к появлению теории вероятностей. В этот период формулируются основные понятия теории: «вероятность», «математическое ожидание», пр.
Развитие и распространение теории вероятностей (XVIII – XIX вв.)	Я. Бернулли К. Ф. Гаусс С. Д. Пуассон	Доказан закон больших чисел. Введен и обоснован «нормальный закон» (закон Гаусса). Доказана центральная предельная теорема. В этот период впервые появляются приложения теории вероятностей.
Появление теории случайных процессов (XX в.)	А. А. Марков	А. А. Марков (ученик П. Л. Чебышева) заложил основы теории «стохастических» случайных процессов
Появление и развитие теории массового обслуживания (XX в.)	А. К. Эрланг А. Я. Хинчин Б. В. Гнеденко	А. К. Эрланг, решая практические задачи совершенствования работы систем связи, вывел ряд формулировок и формул, являющихся основополагающими в теории массового обслуживания. В 30-е гг. XX в. А. Я. Хинчин развивает метод «вложенных цепей Маркова», дающий возможность поиска распределения времени ожидания для простейших потоков на один канал, обслуживающий очередь с произвольным распределением времени обслуживания. В 50-е гг. XX в. Б. В. Гнеденко обобщает формулы Эрланга, рассматривая случаи потери заявок при отказе канала обслуживания и переход заявки на другой свободный канал.

# Основные понятия теории массового обслуживания

Система массового обслуживания (СМО) – математический (абстрактный) объект, содержащий один или несколько приборов  $\Pi$  (каналов), обслуживающих заявки  $З$ , поступающие в систему, и накопитель  $Н$ , в котором находятся заявки, образующие очередь  $О$  и ожидающие обслуживания



*Основными структурными элементами СМО* являются: входящий поток требований, очередь требований, каналы обслуживания и выходящий поток требований.

*Заявка* (требование, запрос, вызов, клиент) – объект, поступающий в СМО и требующий обслуживания в обслуживающем приборе. Совокупность заявок, распределенных во времени, образуют *поток заявок*.

*Обслуживающий прибор* или просто прибор (устройство, канал, линия, пункт, станция) – элемент СМО, функцией которого является обслуживание заявок. В каждый момент времени в приборе на обслуживании может находиться только одна заявка.

*Обслуживание* – задержка заявки на некоторое время в обслуживающем приборе. *Длительность обслуживания* – время задержки (обслуживания) заявки в приборе.

*Накопитель (буфер)* – совокупность мест для ожидания заявок перед обслуживающим прибором. Количество мест для ожидания определяет ёмкость накопителя.

Заявка, поступившая на вход СМО, может находиться в двух состояниях:

- в состоянии обслуживания (в приборе);
- в состоянии ожидания (в накопителе), если все приборы заняты обслуживанием других заявок.

Заявки, находящиеся в накопителе и ожидающие обслуживания, образуют *очередь* заявок. Количество заявок, ожидающих обслуживания в накопителе, определяет *длину очереди*.

*Дисциплина буферизации* – правило занесения поступающих заявок в накопитель (буфер).

*Дисциплина обслуживания* – правило выбора заявок из очереди для обслуживания в приборе.

*Приоритет* – преимущественное право на занесение (в накопитель) или выбор из очереди (для обслуживания в приборе) заявок одного класса по отношению к заявкам других классов.

*Основные предположения при построении математической модели:*

- заявка, поступившая в систему, мгновенно попадает на обслуживание, если прибор свободен;
- в приборе на обслуживании в каждый момент времени может находиться только одна заявка;
- после завершения обслуживания какой-либо заявки в приборе очередная заявка выбирается на обслуживание из очереди мгновенно, то есть, другими словами, прибор не простаивает, если в очереди есть хотя бы одна заявка;
- поступление заявок в СМО и длительности их обслуживания не зависят от того, сколько заявок уже находится в системе, или от каких-либо других факторов;
- длительность обслуживания заявок не зависит от скорости (интенсивности) поступления заявок в систему.

# Классификация систем массового обслуживания (СМО)

Основными признаками классификации СМО являются: количество каналов обслуживания, расположение каналов, возможность образования очереди, дисциплина очереди и объем заявок, обращающихся в системе.

*В зависимости от количества каналов обслуживания:*

- одноканальные СМО;
- многоканальные СМО.

*В зависимости от взаимного расположения каналов:*

- системы с параллельным расположением каналов (в таких системах обслуживание заявок может вести любой свободный канал);
- системы с последовательным расположением каналов (в таких системах обслуживание ведут несколько каналов, причём каждый последующий может приступить к обслуживанию только после того, как обслуживание завершил предыдущий канал).



*В зависимости от возможности образования очереди:*

- системы с отказом в обслуживании - это такие системы, в которых образование очереди невозможно, т.е. если заявка пришла на обслуживания в тот момент, когда все каналы заняты, она получает отказ в обслуживании;
- системы с ограничением на длину очереди - это такие системы, в которых образование очереди будет ограничено каким-либо параметром, например, очередь может иметь ограничения по количеству стоящих в ней заявок (длине очереди) или по времени ожидания (такие СМО называются «системами с нетерпеливыми клиентами»);
- системы без ограничения - в таких системах очередь может расти без ограничения.

*В зависимости от дисциплины очереди:*

- 1) обслуживание «с приоритетом»;
- 2) обслуживание «по правилам».

## Классификация систем массового обслуживания (СМО)



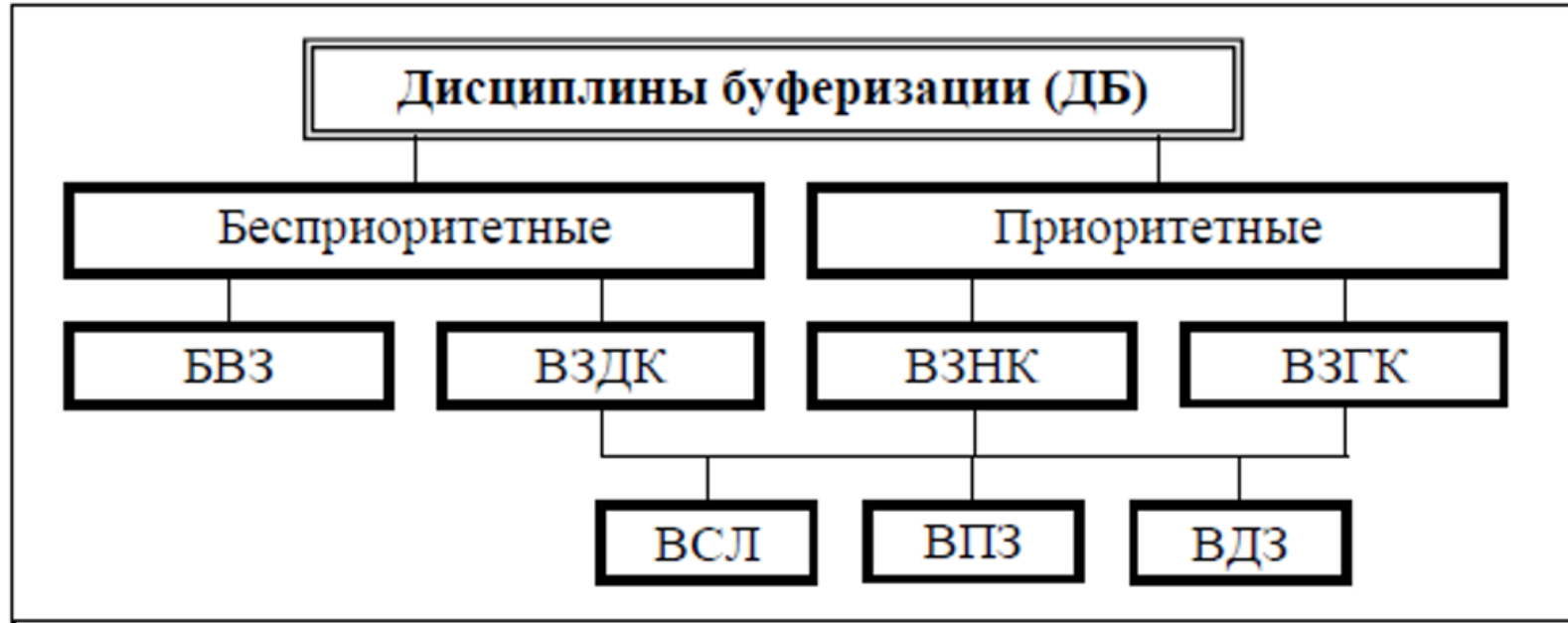
### *Основные виды приоритетов:*

- абсолютный - такой вид приоритета, как правило, имеют надзирательные и контролирующие органы (например, в магазин для проверки приходит санитарная эпидемиологическая станция, магазин закрывается и все остальные заявки (посетители) получают отказ в обслуживании);
- относительный - как правило, имеют работники рассматриваемой системы, например, часто можно наблюдать такую картину: вы стоите в очереди в кассу в крупном супермаркете, подходит продавец, у которого начался обеденный перерыв, и оплачивает свою покупку без ожидания в очереди;
- специальные правила приоритета - закреплены законодательными нормами (законами, постановлениями правительства и/или органами местного самоуправления), в качестве примера можно привести наиболее известное правило специального приоритета: «ветераны ВОВ обслуживаются вне очереди».

### *Основные виды правил:*

- «первый пришел - первый обслужился» - это естественная дисциплина очереди;
- «последний пришел - первый обслужился»;
- обслуживание при случайном отборе заявок из очереди.

## Одна из возможных классификаций дисциплин буферизации



БВЗ - без вытеснения заявок (заявки, заставшие накопитель заполненным до конца, теряются);

ВЗДК - с вытеснением заявки данного класса (то есть такого же класса, что и поступившая);

ВЗНК - с вытеснением заявки самого низкоприоритетного класса;

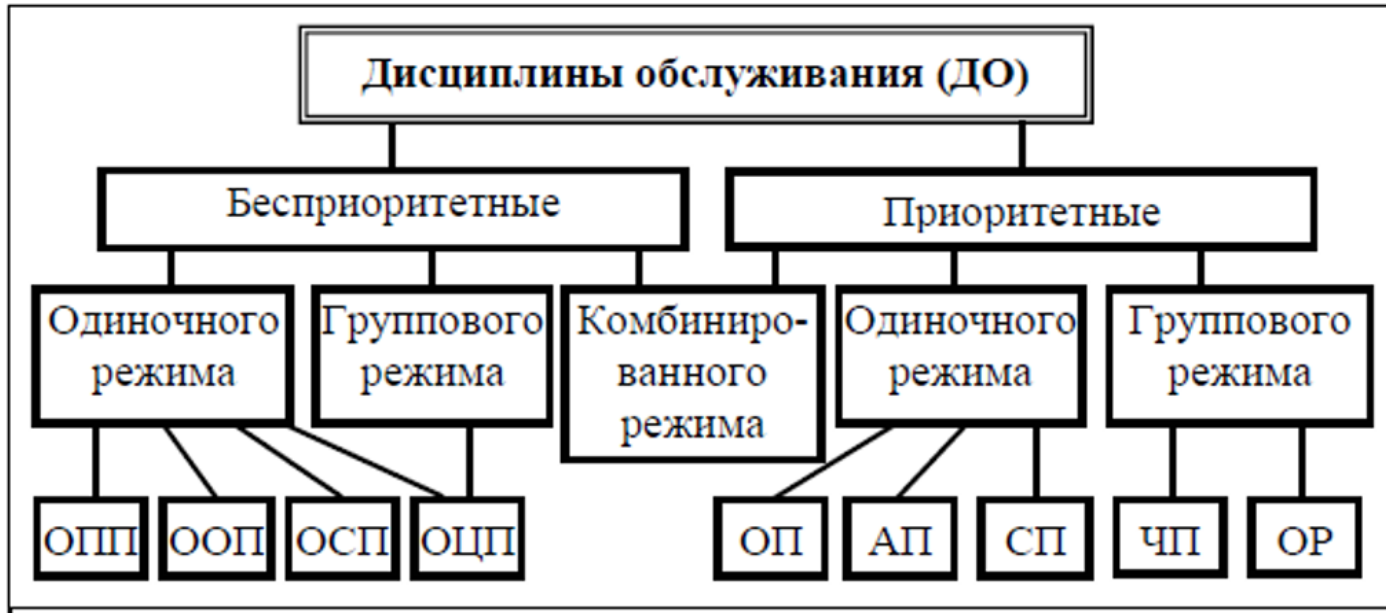
ВЗГК - с вытеснением заявки, принадлежащей группе низкоприоритетных классов.

ВСЛ - вытеснение случайное;

ВПЗ - вытеснение последней заявки (ВПЗ), то есть поступившей в систему позже всех;

ВДЗ - вытеснение «долгой» заявки (ВДЗ), то есть находящейся в накопителе дольше всех.

## Одна из возможных классификаций дисциплин обслуживания



ДО *одиночного режима* – на обслуживание назначается только одна заявка (просмотр очередей с целью назначения на обслуживание в приборе очередной заявки выполняется после обслуживания каждой заявки).

ДО *группового режима* - на обслуживание назначается группа заявок одной очереди.

ДО *комбинированного режима* – комбинация одиночного и группового режимов.

ОПП или FIFO – First In First Out - обслуживание в **порядке** поступления, выбирается заявка, поступившая в систему раньше других;

ООП или LIFO – Last In First Out - обслуживание в **обратном порядке**, выбирается заявка, поступившая в систему позже других;

ОСП - обслуживание в **случайном порядке**, на обслуживание заявка выбирается случайным образом;

ОЦП - обслуживание в **циклическом порядке**, на обслуживание заявки выбираются в процессе циклического опроса накопителей в последовательности 1, 2, ..., N (N – количество накопителей)

ОП - с **относительными приоритетами** (приоритеты учитываются только в моменты завершения обслуживания заявок при выборе новой заявки на обслуживание, поступление в систему заявки с более высоким приоритетом по сравнению с обслуживаемой в приборе не приводит к прерыванию обслуживаемой заявки);

АП - с **абсолютными приоритетами**, означающими, (при поступлении высокоприоритетной заявки обслуживание заявки с низким приоритетом прерывается и на обслуживание принимается поступившая высокоприоритетная заявка);

СП - со **смешанными приоритетами**, представляющими собой любую комбинацию бесприоритетного обслуживания, ОП и АП;

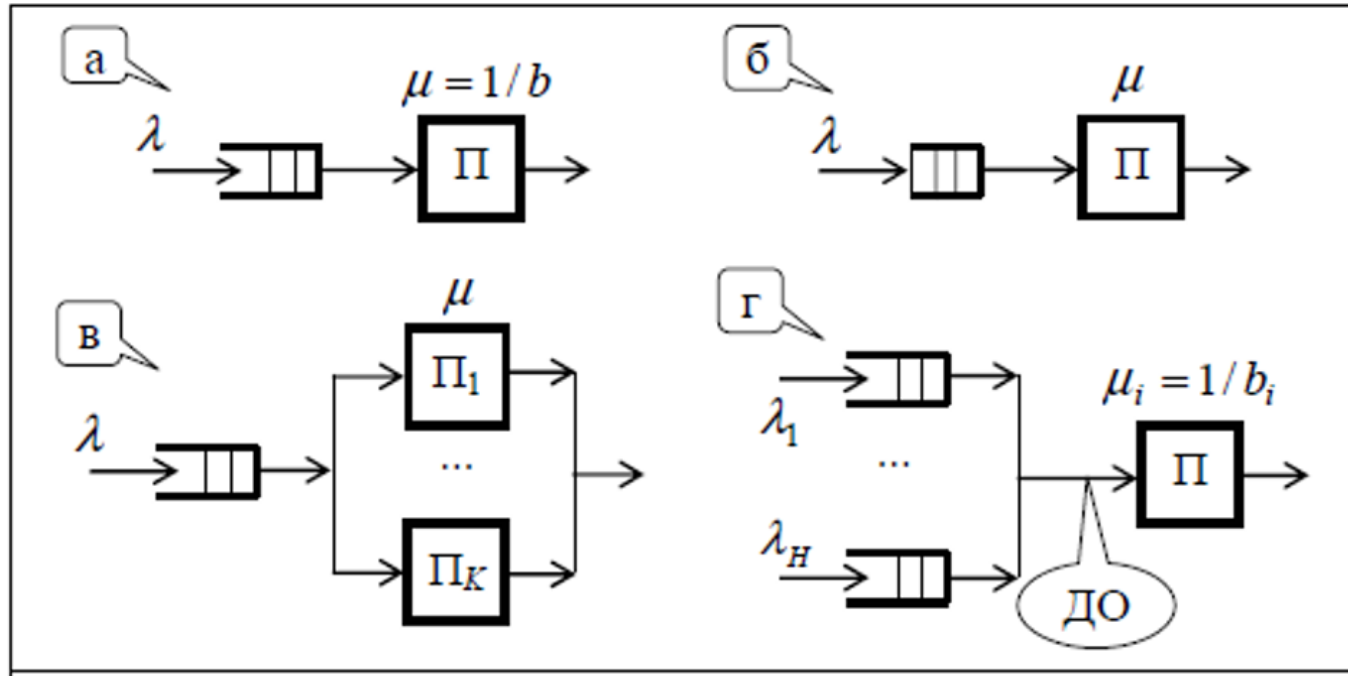
ЧП - с **чередующимися приоритетами**, являющимися аналогом ОП и проявляющимися только в моменты завершения обслуживания группы заявок одной очереди и назначения новой группы;

ОР - обслуживание по **расписанию**, когда заявки разных классов (находящиеся в разных накопителях) выбираются на обслуживание в соответствии с некоторым расписанием (планом).

Аналитическое исследование СМО является наиболее простым, если все потоки событий, переводящие ее из состояния в состояние, — простейшие (стационарные пуассоновские). Это значит, что интервалы времени между событиями в потоках имеют показательное распределение с параметром, равным интенсивности соответствующего потока. Для СМО это допущение означает, что как поток заявок, так и поток обслуживания — простейшие. Под потоком обслуживания понимается поток заявок, обслуживаемых одна за другой одним непрерывно занятым каналом. Этот поток оказывается простейшим, только если время обслуживания заявки  $T_{обсл}$  представляет собой случайную величину, имеющую показательное распределение. Параметр этого распределения  $\mu$  есть величина, обратная среднему времени обслуживания:  $\mu = 1 / b$ , где  $b = M [T_{обсл}]$ . Вместо «поток обслуживания — простейший» часто говорят «время обслуживания показательное».

Если все потоки событий простейшие, то процесс, протекающий в СМО, представляет собой марковский случайный процесс с дискретными состояниями и непрерывным временем. При выполнении некоторых условий для этого процесса существует финальный стационарный режим, при котором как вероятности состояний, так и другие характеристики процесса не зависят от времени.

## Пример классификации базовых СМО



а) одноканальная СМО с накопителем неограниченной ёмкости (СМО без потерь) и однородным потоком заявок;

б) одноканальная СМО с накопителем ограниченной ёмкости (СМО с потерями) и однородным потоком заявок;

в) многоканальная СМО с накопителем неограниченной ёмкости и однородным потоком заявок;

г) одноканальная СМО с накопителями неограниченной ёмкости и неоднородным потоком заявок.

Однородный поток заявок образуют заявки одного класса, а неоднородный поток представляет собой поток заявок нескольких классов. В СМО заявки относятся к разным классам в том случае, если они в моделируемой реальной системе различаются *длительностью обслуживания или/и приоритетами*.

Примечание.

Перечень рассмотренных классификаций не являются исчерпывающим.

Например, *по месту нахождения источника* требования СМО делятся на *замкнутые* и *разомкнутые (открытые)*. В замкнутой системе источник требований находится в самой системе, а в разомкнутой - вне системы.

*По количеству этапов* обслуживания СМО делятся на *однофазные* и *многофазные*. Однофазные СМО – это однородные системы, которые выполняют одну и ту же операцию обслуживания. Многофазные СМО – это системы, в которых каналы обслуживания расположены последовательно и выполняют различные операции обслуживания.

Вывод.

Для построения математической модели СМО необходимо рассмотреть различные модели входящего и выходящего потока требований и процессы, протекающие в системе в процессе ее функционирования.