



КАФЕДРА  
ТЕЛЕМАТИКА

Санкт-Петербургский  
Государственный  
Политехнический  
Университет

Институт прикладной  
математики и механики

**История и методология математики и компьютерных наук**

Тема 3.

**Методология компьютерных наук.**

**Лекция 9 Информационная модель  
ассинхронной, активной платформы вычислений**

---

3 ноября 2022 г.

## Что обсуждали на прошлой лекции

- Взаимодействие большого количества объектов ( частиц) нельзя описать (вычислить траектории) методами механики. **Требуются другая методология.**
- Понятие вероятности, из которого следует **информационная модель мира**, как раз и описывает фундаментальное (базовое) свойств материальной реальности - невозможность детерминированной, т.е. причинно-следственной, характеризации явлений природы.
- Вероятностное описание явлений требует введения еще одного базового понятия (в геометрии это: точка, прямая... в физике – масса, энергия... ) – события. Мера события – это вероятность.
  - в вероятностной аксиоматике Колмогорова имеем определение: **«вероятность — это ограниченная мера  $P(\Omega) = 1$ ».**
  - **в термодинамике - это отношение** количество микросостояний, с помощью которых можно «собрать» конкретное макросостояние, к общему числу возможных микросостояний
- Базовое понятие теории вероятности Колмогорова - это абстрактное «события», которое есть носитель алгебры операций (изоморфной алгебре множеств).
  - Однако, в квантовой механике иная алгебра событий (не алгебра множеств), а «квантовая вероятность» строится отлично от Колмогоровской, поэтому ей соответствует **другая «информационная модель мира».**

## Было отмечено:

- В термодинамике используется не заимствованное из механики понятие скорость частицы, а функция распределения вероятности значения скорости частицы. (средний квадрат скорости молекул :  $\langle v^2 \rangle = 3kT/m$ , ( $v$  около 500 м/с)/
- Теория Колмогорова базируется на аксиомах:
  - Аксиома 1.  $A$  – является **алгеброй событий**, операции которой аналогичны операциям над подмножествами  $\Omega$ .
  - Аксиома 2. Каждому событию  $x$  из  $A$  сопоставляется неотрицательное вещественное число  $P(x)$  – **мера**, которая называется вероятностью,
  - Аксиома 3. Имеется множество элементарных событий  $\Omega$ , причем  $P(\Omega) = 1$  - т.е. **пространство элементарных событий нормировано**
  - Аксиома 4. Если подмножества  $A$ , соответствующие **независимым событиям**  $x$  и  $y$  не пересекаются, то  $P(x+y) = P(x) + P(y)$

Совокупность  $\{ \Omega, A, P \}$  – это **вероятностное пространство** ( поле вероятностей или **носитель информационной меры событий**)

# Существует фундаментальное отличие между «незнанием» и «вероятностью»

4

Ключевое положение: От того, что **вы «чего-то не знаете»**, совсем не значит, что **«это что-то» случайное**.

- Так, если вы не знаете решение уравнения — это не значит, что его решением с одинаковой вероятностью может быть любое число.
- Вероятность можно определить
  - как **степень субъективной уверенности** в истинности суждения, опираясь на теорему Байеса (1702-1761).
  - частота повторения события в длинной серии наблюдений
- **Незнанию** можно сопоставить
  - принцип **максимума энтропии**, который утверждает, что наиболее **достоверным** распределениями вероятностей состояний неопределенной среды с **уверенностью** можно выбрать такое распределения, которые максимизируют меру неопределенности (энтропию) **«поведении» среды**.

# И из этого следует

- Моделирование мира с помощью **вероятностного пространства событий** позволяет определить процесс «извлечения» и «использования» информации из окружающей среды. Информация может рассматриваться как функция и как атрибут.
- Причем «добывание информации» может носить характер «измерения» (активного воздействия) и восприятия (пассивного наблюдения), а процесс передачи данных - носителей информации от «передатчика к приемнику» может иметь как материальный, так и кодовый (символьный) характер. В последнем случае информация отделима от материального носителя. Сложные системы используют обе формы передачи информации
  - Так, биологическая клетка сама «производит» компоненты клеточной мембраны, получая информацию из окружающей среды, которую использует материал для своего «строительства» **Это т.н. процесс аутопоэзиса - основы процессов самоорганизации и, следовательно, перехода материи в новое фазовое состояние, которое характеризуется целевой активностью (жизнь – как активное состояние материи).**

Что из всего этого следует с точки зрения методологии компьютерных наук ?

# Методология компьютерных наук

- компьютерные науки, — это наука о методах и процессах сбора, хранения, обработки, передачи, анализа и оценки **информации** с использованием компьютерных технологий.
  - **Объект** КН – информация и ее представление в целом
  - **Предмет** КН – компьютерные методы и технологии взаимодействия с информацией
- В сферу КН включаются такие разделы как:
  - «наука о данных», программная инженерия, численные методы, дискретная математика, комбинаторика, теория алгоритмов, архитектура компьютерных систем, теория информации, теория кодирования, компьютерные сети, теория языков программирования, криптография, методы машинного обучения и искусственного интеллекта....

# Какие изменения произошли в технологиях вычислительных процессов

За последнее время появились

- Экзоплодные суперкомпьютеры, виртуальные «облачные», «туманные», мобильные, и даже социальные вычислительные системы.
- Их использование породило беспрецедентное количество оцифрованных ресурсов, и появление вычислительных мощностей, создающих общую информационную структуру, которая оставила глубокий «углеродный след от цифровизации»
- Очевидно, что быстрое увеличение разнообразия и объема как вычислительных устройств и цифрового контента создает новые проблемы, ведущие к цифровому хаосу, когда поиск необходимых данных связан с т.н. «информационной перегрузкой» человека так как информация может быть разбросана по различным устройствам и доступ к ней имеют несколько пользователей.

# Динамика развития суперкомпьютерных технологий: ТОР 500

год	число ядер	R <sub>реак</sub> , ПФлопс	R <sub>max</sub> , ПФлопс	эл. мощность, МВт
2022 Frontier	8,700,XXX	1680.XX	1100.XX	21
2020 Fugaku	7,300,XXX	513.XX	415.XX	28
2010 Tianhe-1	186,XXX	4.7X	2.6X	4
2000 ASCI Intel	9,6XX	0.03	0.02	-

- 1 кг угля -> 3 кВтч =0.003 МВтч
  - 1 тонна угля -> 3 МВтч
- 21 МВт -> 21/3 =
- 7 тонн в час
  - 168 тонн в день
  - 60480 тонн в год

Вся Европа потребляет  
400  
000 МВт.

По линии электропередачи (500  
кВ) можно передать

500 М  
Вт

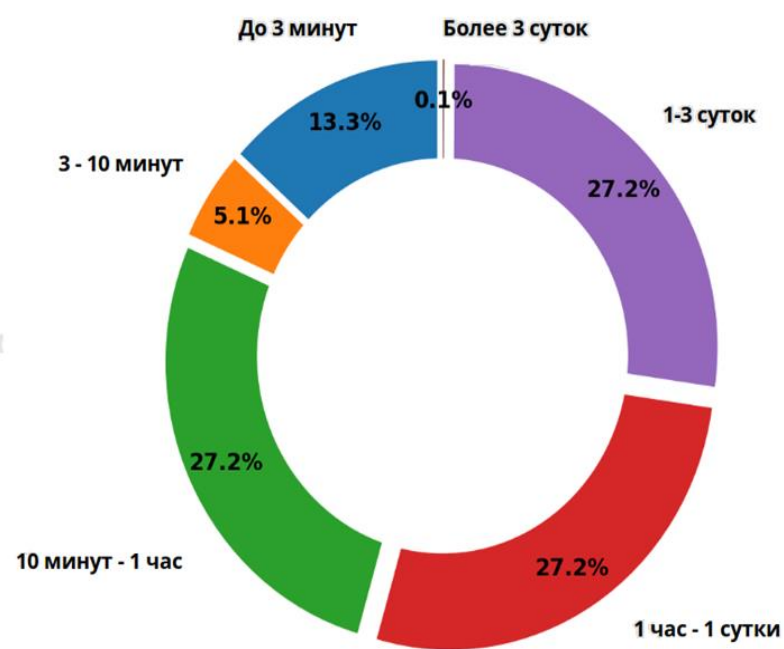
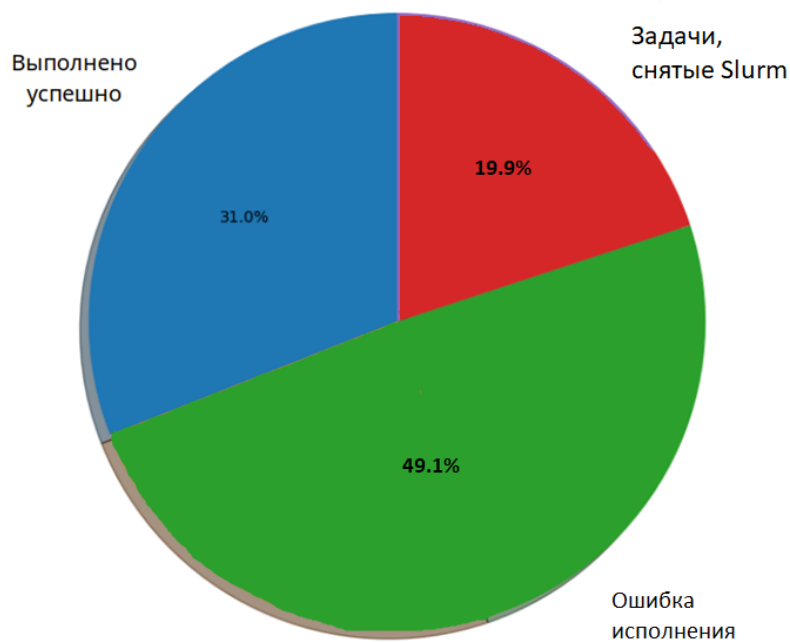
*Проблемы:*

*УПРАВЛЕНИЕ ПРОЦЕССАМИ ПЛАНИРОВАНИЯ ЗАДАНИЙ*

*ПОТРЕБЛЯЕМАЯ МОЩНОСТЬ («углеродный след цифровизации»)*

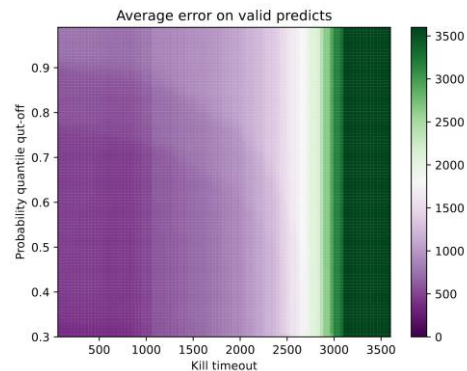


# Мотивация перехода к новым решениям. Статистика «выживания» заданий с СКЦ «Политехнический»



Успешно выполненные составляют **меньше 1/3** от общего числа обработанных заявок пользователей

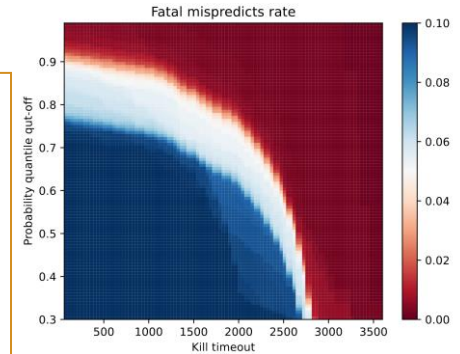
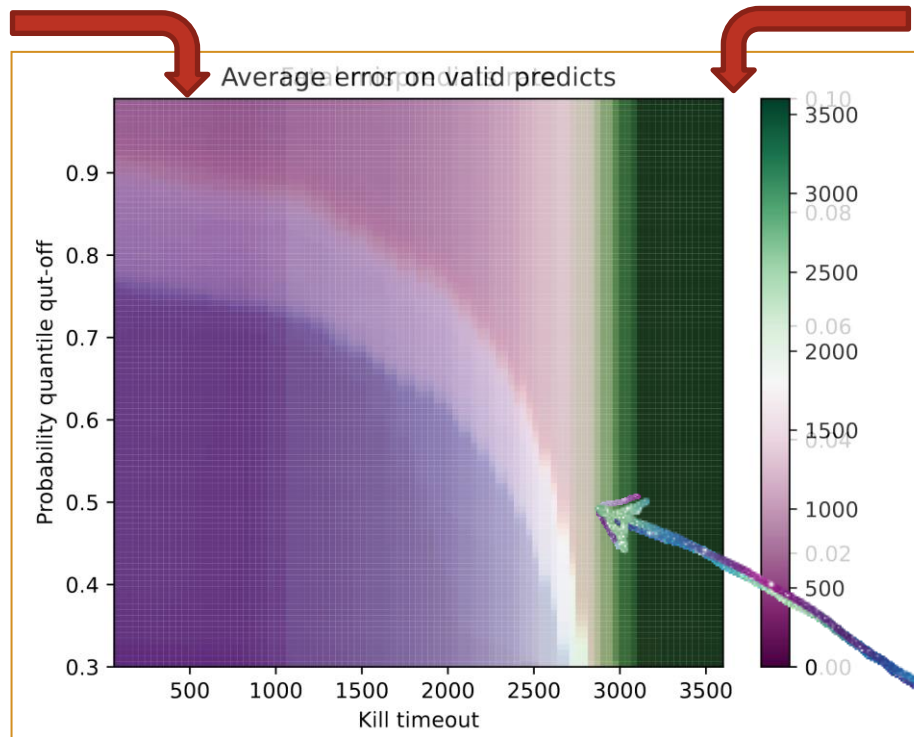
# «устойчивое неравновесие» ошибок вычислений



Комментарий:

Слева – средняя ошибка «**правильных**» предсказаний (время равно или больше фактического)

Справа – доля «**неправильных**» предсказаний (с занижением времени выполнения)

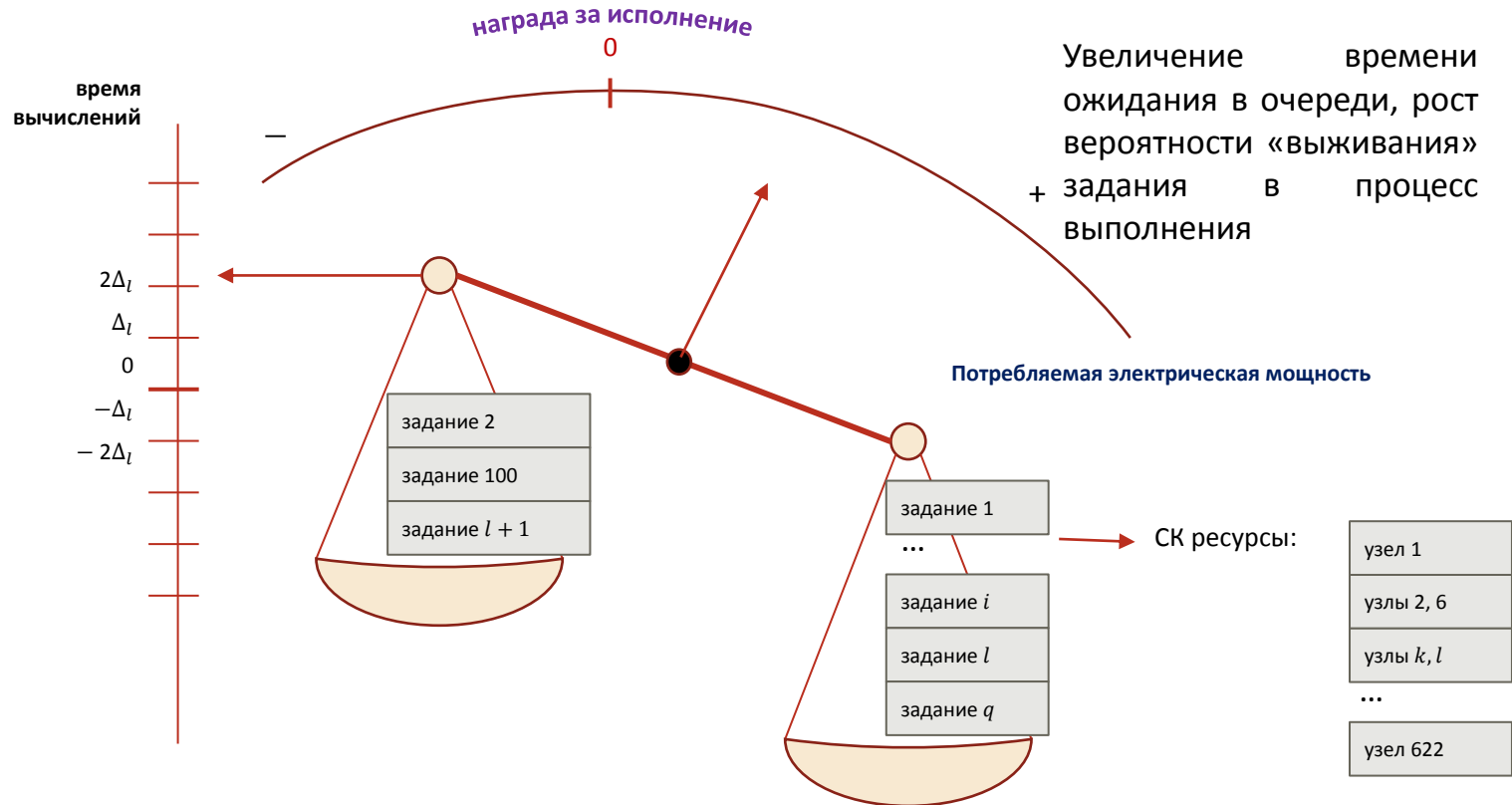


Комментарий:

В качестве критерия «выживаемости» заданий нужна функция «засорения» множества «правильных» множеством «неправильных» ошибок

Такая область на среднем рисунке имеет «розовый цвет» (выделил красным контуром)

Требуется динамическая корректировка параметров приоритета заданий.  
Какие пути решения этой проблемы можно предложить ?



Алгоритмы балансирования нагрузки могут меняться в зависимости от целевых критериев

- 1)  $\min$  среднего времени «простоя» (ожидания)
- 2)  $\min$  «время выполнения» задания
- 3)  $\min$  время ожидания + время выполнения

# Где надо искать решение

- **Асинхронные vs синхронные вычислители** - возможны ли компьютеры без процессоров.
  - Все компьютеры со времен их создания состояли из **процессора, памяти и периферийных устройств**. Причем «активной» частью компьютера является процессор, а память и периферийные устройства только подчиняются командам, которые формирует в соответствии с алгоритмом сам процессор
    - Процессор выставляет на шину адресов адрес необходимой ячейки памяти или порта внешнего устройства.
    - Внешнее устройство или память должны выставить на шине данных свою **информацию**, которую запросил процессор.
    - Процессор, считав информацию с шины данных, обрабатывает ее, записывая в память или во внешнее устройство. **Суть управления компьютером - происходит синхронизация процессора с другими элементами компьютера**. Процесс обмена информации происходит по тактам синхронизации – перехода из одного стабильного состояния в другое.
    - Чем выше тактовая частота синхронизации, тем сложнее согласовывать работу процессора и окружения. Для согласования скорости работы памяти и процессора используются сложные схемы предсказания «переход», кэширование информации и пр. **Все это следствие адресной организации работы памяти вычислителя**.

# Можно ли отказаться от адресной организации памяти ?

- Вполне. Для этого достаточно использовать другой способ организации вычислений, когда и **данные и программы** представляют собой последовательность байтов с заголовком **или пакет**
  - такая вычислительная система должна состоять минимум из четырех основных элементов:
    - коммутаторов, которые доставляют пакеты «получателю»;
    - вычислительных элементов, преобразующих пакеты по определенным правилам - протоколам;
    - элементов памяти, которые хранят пакеты и дают к ним доступ выдают по требованию;
    - периферийных устройств, организующих интерфейс с «внешним миром».
- Действующий прототип такой системы – сеть Интернет

# Сетевая организация вычислений

- Не требует синхронизации всех элементов системы и распределяет функции «центрального процессора» между несколькими вычислительными устройствами – узлами сети и блоками памяти.
- Каждый отдельный узел сети выполняет лишь ограниченный набор функций и поэтому может иметь простую внутреннюю структуру и, следовательно, работать быстро без синхронизации со всеми остальными блоками (синхронизация выполняется через саму коммутируемую среду – в соответствии с дисциплиной обслуживания «задание получите в полете»).

# Компоненты асинхронной вычислительной системы

- Коммутаторы. Основная задача коммутаторов - доставить пакет получателю. Причем в случае, если получателей несколько, коммутаторы должны доставить пакет ближайшему из них.
  - Механизмы поиска оптимального (самого «короткого», самого «надежного», самого «безопасного», ... изохронного и др. ) маршрута известны и отработаны в компьютерных сетях. Любой узел сети, передавая пакет в коммутируемую среду, должен указать только один параметр - адрес получателя. Никакой дополнительной информации не требуется.
- Каналы связи. Для их организации нужно, чтобы коммутаторы «существовали» в общем информационном пространстве . Поэтому объединить два отдельных вычислительных блока в единое целое можно уже с помощью унифицированной системы адресации отдельных узлов (блоков) системы. Коммутатор должен доставить пакет ближайшему свободному блоку, поэтому все одинаковые вычислительные блоки могут иметь один «логический» адрес.

# Организация памяти и процессов взаимодействия АВС

- В отличии от вычислителей, даже одинаковые элементы памяти могут содержать разную информацию, и поэтому должны различаться.
- Адреса портов коммутаторов и узлов памяти могут меняться динамически самой коммутируемой средой, поэтому исчезает проблема добавления нового блока памяти или внешнего устройства.
- Итак, асинхронная вычислительная система становится хорошо масштабируемой, и реальным ограничением ее «размера» служит только разрядность логического адреса, назначаемого отдельным узлам. Наиболее быстрыми каналами являются «симплексы» – однонаправленные каналы передачи данных, но... при этом возрастает сложность физической реализации самой системы.



# Операции в памяти АВС

- Чтобы блоки памяти выполняли свои функции, они должны поддерживать два основных типа операций:
  - ввод в среду коммутации запрашиваемого пакета
  - запись поступающего из коммутируемой среды пакета в блок памяти.
- Ввод пакета – это «запуск» некоторой программы вычислений.
  - в блок памяти поступает пакет-запрос с указанием имени востребованного пакета и набором его аргументов. Блок памяти «ищет» (или формирует) указанный пакет, подставляет в него аргументы и посылает сформированный пакет в коммутируемую среду. Причем одним пакетом-запросом может быть инициировано несколько различных пакетов-ответов. Одним из аргументов может быть адрес получателя, но сам блок памяти "не знает", чем является инициируемый пакет - программой или данными.

**Pro:**

- модульность вычислительной структуры, что позволяет эффективно согласовать возможности аппаратного и программного обеспечения.
- ориентация на многопоточковые вычисления позволяет достигнуть небывалой для синхронных вычислений производительности
- низкое потребление электроэнергии в неиспользуемом узле

**Contra :**

- Ограниченная функциональность вычислительных блоков
- Необходимость модификации модульных процессорных блоков при изменении требований к процессам вычислений.

# Синергия асинхронности и активности

- Activity-Centric Computing (ACC) или вычисления, ориентированные на деятельность позволяют
  - решить глубоко укоренившиеся проблемы управления информацией в традиционных
    - ориентированных на приложения вычислений путем создания унифицированной вычислительной модели, ориентированной на определенные цели "деятельности«, которые специфичны для конкретной области.

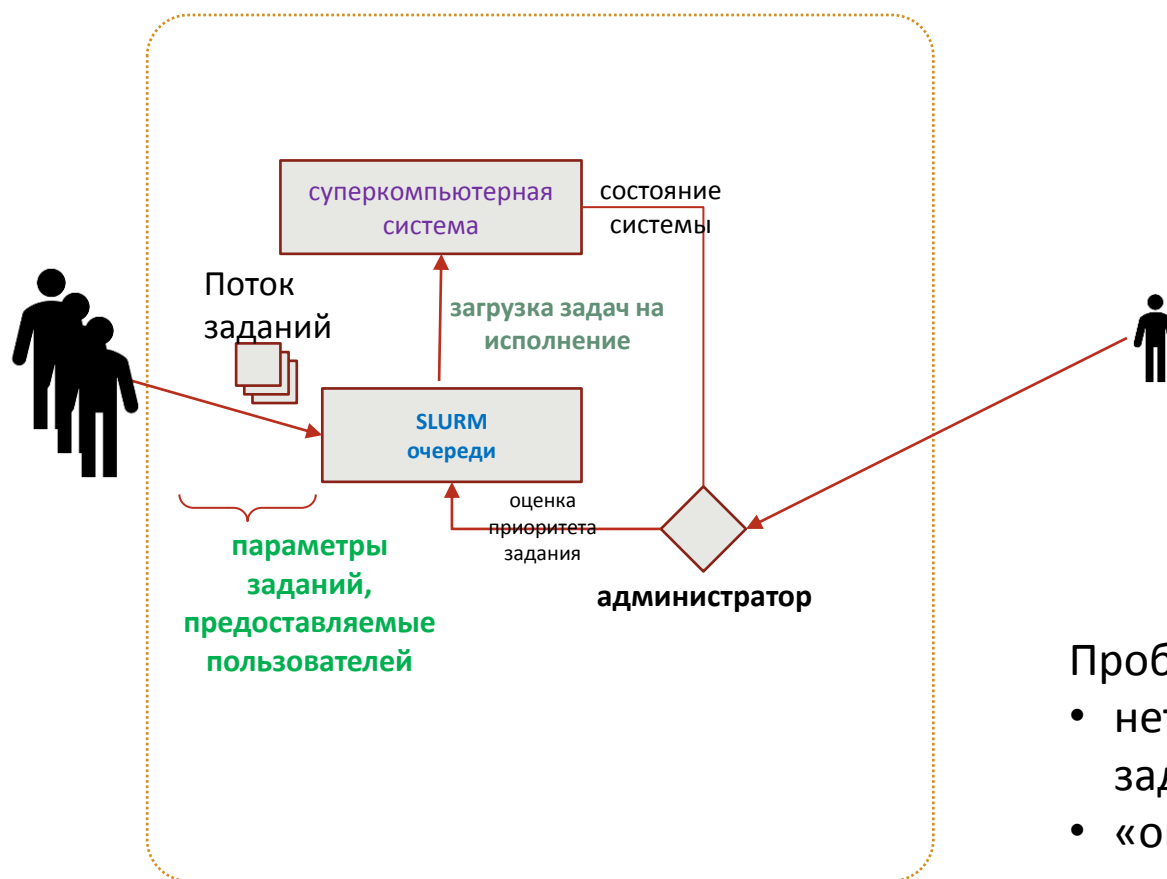
# Понятие активная вычислительная система

Итак, между любыми объектами, сложными системами и организмами могут возникать **взаимодействия различного уровня сложности**:

- первого – вызваны единичными **причинными** связями,
  - второго – вызваны единичными **случайными** воздействиями,
  - третьего – вызваны **постоянно действующими** факторами
- На уровне третьего уровня сложности формируются эпи-процессы эволюции и проходит каскад ко-онтогенезов, что приводит к образованию уже не физико-химических, а информационных связей. Механизм действия таких связей направлен на рост вероятности выживания отдельных особей в рассматриваемой среде.
  - В результате любая **коммуникация** может рассматриваться как информационный механизм формирования «координированного» поведения, которое обеспечивает некое «преимущество» пребывания в социально-ориентированном сообществе
  - В социальных **коммуникациях** нет «передачи информации» в смысле Шеннона: так как сообщения формируются и передаются внутри сообщества, и не используются для связи с внешними объектами – получатель и отправитель сигнала сам объект;
  - Итак, «феномен коммуникации» и использования информационного взаимодействия зависит не от того, **что передаётся**, а от того, как при получении сообщения изменяется состояние «приемника» этого сообщения.

# Как процесс обработки заданий организован сейчас

21

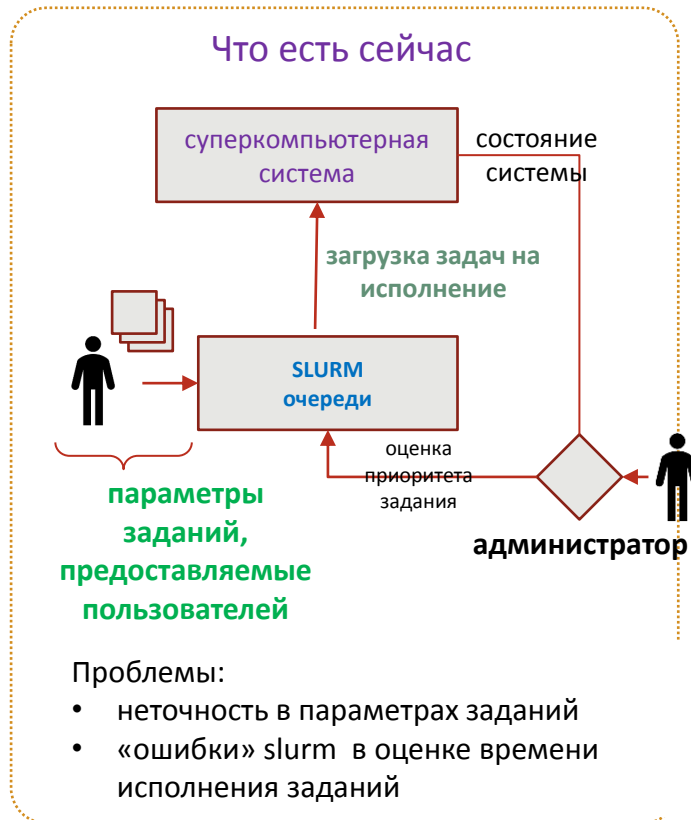


## Проблемы:

- неточность в параметрах заданий пользователей
- «ошибки» slurm в оценке времени исполнения заданий

# Объединения процессов программных вычислений и машинного обучения.

22





## Блок операций вычислений «ВХОД – ВЫХОД»

- Вычислительные блоки АВС имеют один вход ( много входов) и один выход (много выходов). Фактически вычислительный блок работает как **фильтр**, который преобразовывает входящие пакеты по определенным правилам, а результат передают снова в коммутационную среду..
- Набор команд вычислителя может быть ограничен, чтобы гарантировать простую внутреннюю структуру и большую скорость работы. В принципе каждую операцию можно реализовать в виде отдельного узла-устройства, что может существенно ускорить вычисления, но если операций будет слишком много, то могут возникать конфликтные ситуации в самой коммутируемой среде.



# Подходы к формализации : от теории множеств к теории категорий



**Проблема** – «проклятие размерности»



**Проблема** – регуляризация решений обратных задач в базисе функторов

# Заключение

- Сочетание классического сложения вероятностей различных **альтернатив** с классическим выбором одного из нескольких равновероятных путей приводит к «волновому» правилу сложения амплитуд вероятностей.
- Тензорное произведение амплитуд вероятностей определяет принципиально новое состояние – суперпозицию физически несовместных, но информационно возможных состояний объекта.
- Согласно «закону исключенного третьего» одновременное пребывание в несовместных состояний в мире физической реальности запрещено. Суперпозиция возможных состояний объектов является информационной, а не физической характеристикой объектов наделенных памятью.
- Информационная мера вычисляется как функционал от потенциально возможных, но не состоявшихся состояний объекта.