

Распределённые алгоритмы

mk.cs.msu.ru → Лекционные курсы → Распределённые алгоритмы

Блок 2

Вступление:

несколько слов о распределённых системах,
проблемы организации их вычислений,
особенности распределённых алгоритмов

Лектор:

Подымов Владислав Васильевич

E-mail:

valdus@yandex.ru

Распределённые системы

Основные потребности, двигающие развитием распределённых систем:

- ▶ Обмен информацией
- ▶ Совместное использование ресурсов
- ▶ Повышение надёжности за счёт дублирования
- ▶ Повышение производительности за счёт распределения задач
- ▶ Упрощение проектирования систем за счёт специализации
- ▶

Распределённые системы

Примеры распределённых систем:

- ▶ Вычислительные сети
 - ▶ локальные (LAN),
 - ▶ глобальные (WAN)
 - ▶ и другие
(оставим в стороне PAN, MAN и прочие _AN, посчитав их в дальнейшем схожими с LAN и WAN)
- ▶ Многопроцессорные и многоядерные компьютеры
- ▶ Многонитевые программы
- ▶ Системы облачных вычислений
- ▶

Распределённые системы: вычислительные сети

Вычислительная сеть — это совокупность устройств, связанных между собой коммуникационной подсистемой, позволяющей устройствам обмениваться информацией, обычно за счёт **отправки** и **приёма сообщений** (пакетов данных)

Упомянутые примеры типов сетей (LAN, WAN) относятся к их классификации по территориальному охвату:

- ▶ LAN объединяет устройства не очень большой удалённости, отвечающей квартире, зданию, офису, близко расположенным зданиям одной компании, и как правило,
 - ▶ узлы LAN — это небольшие специализированные станции
 - ▶ LAN используется для обмена информацией между знакомыми субъектами и совместного использования ресурсов
- ▶ WAN объединяет устройства разных владельцев и организаций произвольной удалённости, и как правило,
 - ▶ узлы WAN — это разнородные «полноценные» вычислительные устройства
 - ▶ WAN используется для пересылки информации между незнакомыми субъектами на большие расстояния

Распределённые системы: вычислительные сети

Примеры основных различий между LAN и WAN, влияющие на характер распределённых алгоритмов, используемых при их разработке:

▶ Параметры надёжности

- ▶ LAN: узлы можно считать надёжными, и ошибки в их работе — нечастыми исключительными ситуациями
- ▶ WAN: при разработке алгоритмов нельзя предполагать надёжность узлов

▶ Время связи

- ▶ WAN: пересылка сообщения занимает существенно больше времени, чем подготовка к отправке и постобработка после приёма
- ▶ LAN: подготовкой и постобработкой пренебречь нельзя

▶ Однородность

- ▶ LAN: можно достичь «однородности» устройств — согласия о том, какие используются форматы, протоколы, интерфейсы, технологии
- ▶ WAN: разнообразие протоколов, программ, устройств, форматов и т.п. принципиально неустранимо

Распределённые системы: вычислительные сети

Примеры основных различий между LAN и WAN, влияющие на характер распределённых алгоритмов, используемых при их разработке:

▶ Взаимное доверие

- ▶ LAN: устройства по умолчанию полагаются заслуживающими доверия как имеющие общего контролирующего владельца
- ▶ WAN: следует полагать, что среди узлов сети есть враждебные, и обеспечивать защиту от них

▶ Способ соединения

- ▶ WAN: сеть представляет собой набор узлов, соединённых каналами типа точка-точка
- ▶ LAN: используются разные виды соединений, в том числе широковещательные, включая соединение через общую шину

Распределённые системы: вычислительные сети

Особенностями построения сетей определяется и спектр возникающих алгоритмических проблем — например:

- ▶ Обеспечение широкого вещания и синхронизации (LAN) и доставки сообщений адресату (LAN, WAN)
- ▶ Маршрутизация пакетов — выбор путей и каналов для доставки (WAN)
- ▶ Балансировка нагрузки узлов и каналов
- ▶ Предотвращение блокировки при передаче данных
- ▶ Обеспечение безопасности узлов (WAN)
- ▶ Эффективное избрание лидера (LAN)
- ▶ Обнаружение завершения вычислений (LAN)
- ▶ Поиск и назначение ресурсов, распределение с взаимным исключением (LAN)

Распределённые системы: многопроцессорные компьютеры

Многопроцессорный компьютер, как видно из названия, — это вычислительное устройство, в одном корпусе которого размещено несколько (более одного) процессоров

Компьютер с такой архитектурой может использоваться для ускорения вычислений (тогда это **параллельный компьютер**) или для повышения надёжности вычислений (тогда это компьютер **с дублированием**)

В параллельном компьютере требуемое вычисление разбивается на подвычисления, выполняющиеся на разных узлах-процессорах

В компьютере с дублированием вычисление воспроизводится на нескольких узлах, после чего результаты совмещаются с исправлением возможных ошибок

Распределённые системы: многопроцессорные компьютеры

Многопроцессорный компьютер можно расценивать как крайне локализованную и крайне однородную LAN, узлы которой совместно решают единую вычислительную задачу

Особенностями устройства многопроцессорных компьютеров определяются и особые по сравнению с вычислительными сетями алгоритмические проблемы — например:

- ▶ Реализация коммуникационной подсистемы в «хорошей» топологии сети
- ▶ Обеспечение доступа к разделяемой памяти
- ▶ Устойчивость к отказам доверенных узлов

Распределённые системы: взаимодействующие процессы

Разработку сложной программной системы можно упростить, если проектировать её как (распределённую) совокупность процессов, каждый из которых решает свою простую задачу

Даже если процессы физически будут выполняться на одном вычислительном устройстве, они могут рассматриваться как **логически обособленные** и взаимодействующие по строго заданным каналам связи, то есть тоже как полноценная распределённая система

С таким разделением программы на обособленные процессы также связаны алгоритмические проблемы, характерные для распределённых систем — например:

- ▶ Обеспечение атомарности операций чтения и записи
- ▶ Синхронизация процессов и обмен сообщениями
- ▶ Сборка мусора

Архитектура сети

Коммуникационная подсистема в современных распределённых системах строится из модулей, каждый из которых выполняет свою узкую специальную функцию и взаимодействует с другими модулями при помощи строго определённых узкоспециализированных интерфейсов

Архитектурой сети называется совокупность этих модулей с определениями соответствующих **интерфейсов** и протоколов взаимодействия

В современных архитектурах модули выстраиваются в **иерархию уровней** абстракции взаимодействия, и каждый уровень (k -й) может обращаться только к предшествующим ему в иерархии ($(k - 1)$ -му) посредством соответствующего интерфейса ($(k - 1/k)$)



Архитектура сети

В связи с бурным развитием и активным использованием сетевых технологий (*уже очень и очень давно*) появилась необходимость в стандартизации архитектуры

В рамках такой стандартизации возникли и заняли нишу стандартные сетевые архитектуры, такие как

- ▶ ISO OSI
 - ▶ ISO = International Organization for Standardization
 - ▶ OSI = Open-Systems Interconnection
- ▶ IEEE 802
 - ▶ IEEE = Institute of Electrical and Electronics Engineers
 - ▶ 802 — исторически сложившийся бессмысленный порядковый номер
- ▶ IETF Internet protocol suite (TCP/IP)
 - ▶ IETF = Internet Engineering Task Force

Модель OSI

Модель OSI состоит из семи уровней:

7. Прикладной
6. Уровень представления
5. Сеансовый
4. Транспортный
3. Сетевой
2. Канальный
1. Физический

Модель OSI

1. Физический уровень предназначен для передачи последовательностей битов по каналам связи

На этом уровне цифровые биты преобразуются в уровни напряжения, электрические и оптические сигналы, электромагнитные волны и т.п.

2. Канальный уровень предназначен для передачи информации через физический уровень с обнаружением и исправлением ошибок, возникающих из-за физической природы среды передачи

Для этого, как правило, поток битов делится на сообщения одинакового размера: **фреймы** — снабжаемые дополнительной проверочной информацией, и в пересылку фрейма добавляется общение отправителя и получателя посредством служебных сообщений

3. Сетевой уровень предназначен для обеспечения связи между узлами сети, в том числе если они не соединены каналом связи напрямую

На этом уровне определяются **маршруты** пересылки сообщений (**пакетов**) для достижения ими требуемого узла в условиях имеющихся каналов и их загруженности

Модель OSI

4. Транспортный уровень предназначен для обеспечения надёжной адресованной передачи информации через сетевой уровень

Некоторые подходы, применяющиеся на этом уровне, схожи с подходами канального уровня к устранению ошибок при передаче сообщений

5. Сеансовый уровень предназначен для поддержания **сеанса** связи в рамках длительного обмена информацией: **открытие (создание)** и **закрытие (завершение)** сеанса, обмен в рамках сеанса, синхронизация и **восстановление** сеанса при сбое

6. Уровень представления предназначен для стыковки протоколов и форм представления данных в разных узлах: кодирования и декодирования, сжатия и распаковки, шифрования и дешифрования, ...

7. Прикладной уровень предназначен для передачи данных пользовательскими приложениями

Модель OSI

Канальный уровень в **локальных сетях**, допускающих подключение многих узлов к одной физической среде передачи данных, принято (согласно *IEEE 802*) разделять на два подуровня:

2.2. **LLC** (Logical Link Control, **управление логической связью**):
подуровень, взаимодействующий с сетевым и решающий основные задачи канального уровня

2.1. **MAC** (Media Access Control, **управление доступом к среде**):
подуровень, разрешающий конфликты доступа узлов к общей среде передачи данных

Особенности распределённых программ

Язык, использующийся для реализации распределённых систем, должен поддерживать следующие возможности:

- ▶ **Параллелизм**: создание автономно выполняющихся **процессов**

Создание может быть как статическим, так и динамическим, и сами процессы, как правило, представляют собой последовательно выполняющиеся несложные подпрограммы

- ▶ **Коммуникация**: обмен информацией между процессами

Как правило, для этого используется или **система обмена сообщениями**, или **общая (разделяемая)** памяти, специально предназначенная для такого обмена

Особенности распределённых программ

Язык, использующийся для реализации распределённых систем, должен поддерживать следующие возможности:

- ▶ **Недетерминизм**: продолжение вычисления разными способами в зависимости от событий внутри процесса и в его окружении

В частности, недетерминизм вносится

- ▶ приёмом сообщения: сообщение может быть готово для приёма, а может и не быть и
- ▶ отправкой сообщения: среда может быть готова принять сообщение, а может и быть

Типовой способ внесения подобного недетерминизма — это снабжение команд **предусловиями** (по-другому — предохранителями, охранами, защитой; англ. guard) — булевыми выражениями, истинность которых означает готовность команды к выполнению (к выполнению в общем случае могут быть одновременно готовы и несколько команд)

Особенности распределённых алгоритмов

Характерными особенностями устройства распределённых систем определяются и основные особенности алгоритмов, предназначенных для реализации в узлах таких систем:

- ▶ Незнание глобального состояния

Узел распределённой системы

- ▶ располагает своим состоянием и информацией, переданной ему посредством сообщений и
- ▶ неспособен напрямую обозревать состояние других узлов и коммуникационной подсистемы

Поэтому узел принципиально не может использовать понятие глобального состояния системы

Особенности распределённых алгоритмов

Характерными особенностями устройства распределённых систем определяются и основные особенности алгоритмов, предназначенных для реализации в узлах таких систем:

▶ Отсутствие глобального времени

Часы разных узлов могут иметь разные абсолютные значения и отсчитывать время в разном темпе

Если в узле требуется отсчёт времени, то он может руководствоваться только показаниями своих (локальных) часов и информацией о значениях других часов, переданной через сообщения

Поэтому узел принципиально не может использовать понятие времени, единого для всей системы, и должен использовать другие способы отсчёта времени, основанные только на доступной ему информации

Особенности распределённых алгоритмов

Характерными особенностями устройства распределённых систем определяются и основные особенности алгоритмов, предназначенных для реализации в узлах таких систем:

▶ Недетерминизм

Быстродействие компонентов распределённой системы может различаться и изменяться с течением времени в связи с физическими явлениями, выбором аппаратуры и программного обеспечения и сбоями в их работе и другими факторами

Узел должен быть способным корректно продолжить выполнение независимо от времени получения информации от окружения

Отсутствие глобального состояния и глобального времени и недетерминизм в совокупности приводят к тому, что **правильное** решение даже (*казалось бы*) простых задач в распределённой системе оказывается непростым и довольно замысловатым