

Распределенные информационные системы

Правда и ложь о РИС

Fallacies of distributed computing

Вопросы

- История
- Общие описание заблуждений
 - Доставка данных
 - Архитектура [сети]
 - Влияние человеческого фактора
- Последствия заблуждений

ИСТОРИЯ ВОПРОСА

История

- 1994 Peter Deutsch – сотрудник Sun опубликовал список из 7 предположений, которые часто делают архитекторы и проектировщики РИС, и которые оказываются неверными в перспективе, вызывая различные проблемы при эксплуатации РИС
- 1997 James Gosling – добавил еще одно такое предположение
- Общий список называется «The 8 fallacies of distributed computing»

Доставка данных

Архитектура [сети]

Влияние человеческого фактора

ОБЩЕЕ ОПИСАНИЕ ЗАБЛУЖДЕНИЙ

Заблуждения: Доставка данных

- Доставка гарантирована (the network is reliable)
- Задержки доставки нет (latency is zero)
- Пропускная способность не ограничена (bandwidth is infinite)
- Цена доставки нулевая (transport cost is zero)

Заблуждения: Архитектура

- Топология сети неизменна (topology doesn't change)
- Сеть однородна (the network is homogeneous)

Заблуждения: Человеческий фактор

- Сеть безопасна (the network is secure)
- Управление сетью централизовано (there is one administrator)

ПОСЛЕДСТВИЯ ЗАБЛУЖДЕНИЙ

Доставка гарантирована – 1/2

- ПО **обычно** пишется с небольшим учетом того, что сеть [доставки данных] не гарантирует доставку данных:
 - Во время сбоя в сети так устроенные приложения могут, например:
 - Останавливаться
 - Бесконечно ждать ответного пакета
 - После восстановления сети приложения также могут не выполнять повторные попытки или требовать [ручной] перезагрузки

Доставка гарантирована – 2/2

- Более сложные проблемы:
 - Остановка во время выполнения одной из операций требует приостановки операций остальных приложений (включая не только непосредственно взаимодействовавших)
 - Сбой одной из операций требует «отката» всей цепочки операций
 - Взаимодействие с реальными объектами не допускает превышения времени ответа
 - «Непрошедшая» операция приводит к рассогласованию состояний элементов РИС (включая части распределенной БД)

Нет задержки доставки – 1/3

- Недооценка задержки доставки пакетов в сети может проявиться как:
 - Стохастически проявляющиеся сбои при переходе к РИС от реализации алгоритма в монолитной архитектуре или в ЛВС
 - Неверное обнаружение потерь пакетов на транспортном уровне и уровне приложений → увеличивается трафик (из-за неоправданной повторной посылки пакетов) → «съедается» пропускная способность

Нет задержки доставки – 2/3

- Неочевидные моменты:
 - С точки зрения разработчика неверная оценка максимально допустимой задержки проявляется как негарантированность доставки (см. выше), что еще больше затрудняет поиск причин сбоя и отладку РИС
 - Задержки доставки меняются
 - Доступная пропускная способность влияет на текущую задержку доставки
 - Сеть (кроме ЛВС по построению) анизотропна
 - Задержки состоят из: времени установления соединения, собственно доставки данных

Нет задержки доставки – 3/3

- Задержки доставки:
 - Зависят от расстояния
 - Световая секунда = 299 792 458 метров
 - 1 мс ~ всего 300 км!
 - Зависят от маршрута пакета (расстояние + задержки в устройствах)
 - «Прыгают» при переконфигурации сети (редко)
 - Зависят от текущей загрузки устройств (чаще)
 - Зависят от факторов внешней среды:
 - Смена политик качества обслуживания провайдером (даже и не ваших политик)
 - Погода в целом, электрические атмосферные явления, солнечная активность

Пропускная способность не ограничена – 1/2

- Незнание и/или игнорирование ограничений пропускной способности даже частью поставщиков данных:
 - Образование «бутылочных горлышек»
 - Недоставка или задержка доставки данных

Пропускная способность не ограничена – 2/2

- Неочевидные моменты:
 - Перегрузка сети приводит к возрастанию задержек → Задержки доставки данных приводят к превышению таймаутов → Проявляется как негарантированность доставки
 - Скорость доставки данных зависит от их объема
 - 100 Мбит сеть x 1 КБайт → 80 мкс
 - 100 Мбит сеть x 1 МБайт → 80 мс

Цена доставки нулевая – 1/2

- «Скрытые» затраты на построение и поддержки сетей и подсетей заметны даже на фоне общей стоимости проекта и должны быть в явном виде указаны в бюджете:
 - Несоблюдение бюджета
 - Несоответствие характеристик сети требуемым
 - Несогласованности требований к сети со стороны различных пользователей

Цена доставки нулевая – 2/2

- Неочевидные моменты:
 - Отсутствие пункта «построение и поддержание функционирования сети» в бюджете может быть признаком того, что постановка задачи на построение сети и ее проектирование «выпали из рассмотрения»
 - Поддержание функционирования сети требует обычно не только финансовых, но и человеческих ресурсов (+ организационных, правовых, etc.)

Топология сети неизменна – 1/2

- Сами по себе изменения в топологии сети [обычно] «прозрачны» для вышележащих уровней, однако:
 - Вызывают изменение пропускной способности (временное и/или постоянное) → см. выше
 - Вызывают изменение задержек (временное и/или постоянное) → см. выше
 - Могут изменять связность

Топология сети неизменна – 2/2

- Неочевидные моменты:
 - В тестовом (сборочном, отладочном, демонстрационном) окружении невозможно смоделировать в полной мере все возможные комбинации изменений топологии, которые есть (будут) при реальной эксплуатации РИС
 - Изменение топологии сети включает, кроме изменения связей между устройствами:
 - Произвольное исчезновение и восстановление элементов самой РИС
 - Появление новых элементов РИС
 - Совместное использование сети несколькими РИС
 - Пользователи мигрируют по сети → роли хостов меняются

Сеть однородна - 1/2

- Если РИС предполагает, что сеть однородна, то это может привести к тем же проблемам, которые возникают в результате первых трех ошибок (сеть надежна, задержка равна нулю, ширина полосы пропускания бесконечна)

Сеть однородна - 2/2

- Неочевидные моменты:
 - Различные устройства подключаются разными способами:
 - Проводная ЛВС
 - Беспроводная ЛВС
 - Мобильная связь
 - Неоднородность проявляется как:
 - Различия в каналах связи: пропускная способность, физическая организация (кабель, оптоволокно, радиоканалы, etc.)
 - Различное сетевое оборудование
 - Доступные приложениям протоколы

Сеть безопасна – 1/4

- Поддержание сети в безопасном состоянии является непрерывным процессом:
 - Обнаруживаются новые «дыры» в безопасности, которые необходимо закрывать
 - Движение персонала требует действий также и в настройках безопасности
 - Обновление СПО
 - Масштабирование РИС

Сеть безопасна – 2/4

- Неочевидные моменты:
 - По мере устаревания «дыр» в безопасности возможность ими воспользоваться получают не только «суперхакеры»:
 - Появляются готовые «ломалки» → любые злоумышленники
 - Вирусы и трояны
 - Для понимания способов противодействия злоумышленникам надо определить, какой именно вред может быть нанесен:
 - Уничтожение информации
 - Воровство информации (статической)
 - Воровство информации (динамической)
 - Нарушение функционирования РИС

Сеть безопасна – 3/4

- Неочевидные моменты (2):
 - Для значительной части проблем есть готовые решения:
 - Обновления ОС и ОПО
 - Антивирусы и сетевые экраны с обновляемыми базами
 - ~~Люди – это~~ Человеческий фактор:
 - Сознательное нанесение вреда
 - ~~Ребенок незагружен~~ Текущая активность:
 - Бесконтрольная установка «удобного» ПО
 - Использование социальных сетей

Сеть безопасна – 4/4

- Проблема «Неуловимого Джо»:
 - + Часто проблема сетевой безопасности неактуальна, так как РИС злоумышленникам неинтересна
 - Получение контроля над хостами РИС может быть использовано для «внешних» задач:
 - DDOS
 - Промежуточные сервера в цепочке удаленного доступа
 - Хранение незаконного контента
 - Использование ресурсов: запуск [незаконных] сервисов

Управление сетью централизовано – 1/2

- Децентрализованность проявляется несколькими способами:
 - Внутри РИС: подразделения компании имеют своих администраторов
 - Вне РИС: локальная сеть подразделения + местный провайдер + глобальный провайдер
- Создаваемые проблемы:
 - Внутренняя + внешняя: некоторые сервисы недоступны или доступны по-разному (см. выше Сеть однородна)
 - Внешняя: сбои или изменения свойств сети (см. выше Топология сети неизменна)

Управление сетью централизовано – 2/2

- Неочевидные моменты:
 - Не существует коробочных решений для централизованного управления неоднородной сетью
 - В географически распределенных РИС наличие нескольких «центров управления» обусловлено практикой использования РИС (оперативность действий локального администратора выше)
 - Работа 7/24 требует сменности работы администраторов, что вызывает дополнительные трудности с передачей оперативных задач, решений и действий по ним