

Примеры инфраструктурных решений, применяющихся в крупных сетевых проектах

Пример реализации инфраструктуры в Google

Распределенная инфраструктура системы: GFS, MapReduce и BigTable

При этом поддерживаются два принципа:

- Легкое развертывание для компании при низком уровне издержек
- больше денег вкладывается в оборудование для исключения возможности потерь данных

Распределенная файловая система GFS (Google File System)

Файлы в GFS организованы иерархически, при помощи каталогов, с файлами можно выполнять обычные операции: создание, удаление, открытие, закрытие, чтение и запись.

В системе существуют мастер-сервера и чанк-сервера, собственно, хранящие данные. Как правило, GFS-кластер состоит из одной главной машины мастера и множества машин, хранящих фрагменты файлов чанк-серверы. Клиенты имеют доступ ко всем этим машинам. Файлы в GFS разбиваются на куски — чанки.

Каждое изменение чанка должно дублироваться на всех репликах и изменять метаданные. В GFS мастер дает чанк во владение одному из серверов, хранящих этот чанк. Такой сервер называется первичной репликой. Остальные реплики объявляются вторичными.

Мастер важное звено в системе. Он управляет репликациями чанков: принимает решения о размещении, создает новые чанки, а также координирует различную деятельность внутри системы для сохранения чанков полностью реплицированными, балансировки нагрузки на чанксерверы и сборки неиспользуемых ресурсов. В отличие от большинства файловых систем GFS не хранит состав файлов в директории. GFS логически представляет пространство имен, как таблицу, которая отображает каждый путь в метаданные.

Организация работы с данными при помощи MapReduce

MapReduce является программной моделью и соответствующей реализацией обработки и генерации больших наборов данных. Пользователи могут задавать функцию, обрабатывающую пары ключ/значение для генерации промежуточных аналогичных пар, и сокращающую функцию, которая объединяет все промежуточные значения, соответствующие одному и тому же ключу.

Система берет на себя детали разбиения входных данных на части, составления расписания выполнения программ на различных компьютерах, управления ошибками, и организации необходимой коммуникации между компьютерами.

Пример реализации инфраструктуры для проекта Flickr

Входные запросы поступают на сдублированные контроллеры приложений Brocade ServerIron ADX. Они обеспечивают коммутацию приложений и балансировку трафика, основываясь на принципе виртуальных ферм серверов: коммутатор приложений получает все клиентские запросы. Выбор “лучшего” сервера производится на основании механизма Real-Time Health и

наличия требуемой производительности Последовательно повышается коэффициент использования для всех серверов. Интеллектуальное распределение загрузки осуществляется для всех доступных ресурсов Метод конфигурируется и выбирается пользователем. Обеспечивается защита серверной фермы от атак и от неправильной эксплуатации Клиенты подсоединяются к серверам приложений используя виртуальный IP (VIP). VIP адреса настраиваются на коммутаторе приложений Коммутатор приложений осуществляет трансляцию адресов после выбора нужного сервера, причем сами адреса серверов скрыты. Обслуживание сессий ведется согласно последовательности: запись о каждой пользовательской сессии создается в таблице Каждая сессия назначается определенному серверу. Все сообщения в рамках сессии посылаются к одному серверу Таблицы сессий синхронизируются между двумя коммутаторами.