**Частные вопросы управления ИТ-инраструктурой**

Сегодня Google это далеко не только поисковая система. Это огромная сервисно-ориентированная платформа для построения масштабируемых приложений, позволяющая выпускать и поддерживать множество конкурентоспособных интернет-приложений, работающих на уровне всей глобальной сети. Компания ставит перед собой цель постоянно строить все более и более производительную и масштабируемую инфраструктуру для поддержки своих продуктов. Рассмотрим основные принципы и компоненты этого инфраструктурного решения. Но сначала несколько цифр, для того, чтобы представить масштабы решения.

Система BigTable позволяет хранить миллиарды ссылок (URL), сотни терабайт снимков со спутников, а также настройки миллионов пользователей Google визуализирует свою инфраструктуру в виде трехслойного стека:Продукты: поиск, реклама, электронная почта, карты, видео, чат, блоги и т.п. Распределенная инфраструктура системы: GFS, Map Reduce и BigTable Вычислительные платформы: множество компьютеров во множестве датацентров При этом поддерживаются два принципа: Легкое развертывание для компании при низком уровне издержек Больше денег вкладывается в оборудование для исключения возможности потерь данных С точки зрения ИТ-инфраструктуры нас интересуют в первую очередь решения, расположенные на втором слое.

Пример реализации инфраструктуры для проекта FlickrFlickr является мировым лидером среди сайтов размещения фотографий. Перед Flickr стоит крайне непростая задача, они должны контролировать огромное количество ежесекундно обновляющегося контента, непрерывно пополняющиеся пользователи, постоянный поток новых предоставляемых пользователям возможностей, и при этом поддерживать постоянно высокий уровень производительности.123

Статистика Более четырех миллиардов запросов в день Примерно 35 миллионов фотографий в кэше прокси-сервера Squid Около двух миллионов фотографий в оперативной памяти Squid Всего приблизительно 470 миллионов изображений, каждое представлено в 4или 5 размерах38 тысяч запросов к memcached (12 миллионов объектов)2 петабайта дискового пространства Более 400000 фотографий добавляются ежедневно Использующиеся программные компоненты Примечательно, что на проекте Flickr используется практически только свободное программное обеспечение. Платформа GNU/Linux (RedHat)СУБД MySQLWeb-сервер ApacheСкрипты программной логики, написанные на языке PHP и PerlСредства сегментирования (Shards) (прим.: разбиение системы на части, обслуживающие каждая свою группу пользователей; называть можно было по-разному, но давайте остановимся на этом варианте перевода)Memcached для кэширования часто востребованного контентаSquid в качестве обратного прокси-сервера для html-страниц и изображений Шаблонизатор SmartyPEAR для парсинга e-mail и XML ImageMagick для обработки изображений SystemImager для развертывания элементов конфигурации Ganglia для мониторинга распределенных систем Subcon для хранения важных системных конфигурационных файлов в SVN-репозитории для легкого развертывания на машины в кластере Cvsup для распространения и обновления коллекций файлов по сети.