
OPEN SOURCE BUSINESS CONFERENCE
Building Your Big Data Future with Open Source
COMPUTERWORLD
OSBC
SAN FRANCISCO

Big Inbox. Big Data. Big Problems?

Utpal Thakrar
Principal Product Manager
Openwave Systems Inc.

Building Your Big Data Future with Open Source

Openwave is a global software innovator delivering context-aware mediation and messaging solutions that enable mobile operators and the broader ecosystem to create and deliver smarter services.

- Silicon valley based company, since 1996
- Serving tier-1, tier-2 mobile and broadband operators worldwide
- Product Focus - Traffic Mediation, Analytics, Data Mediation, Messaging

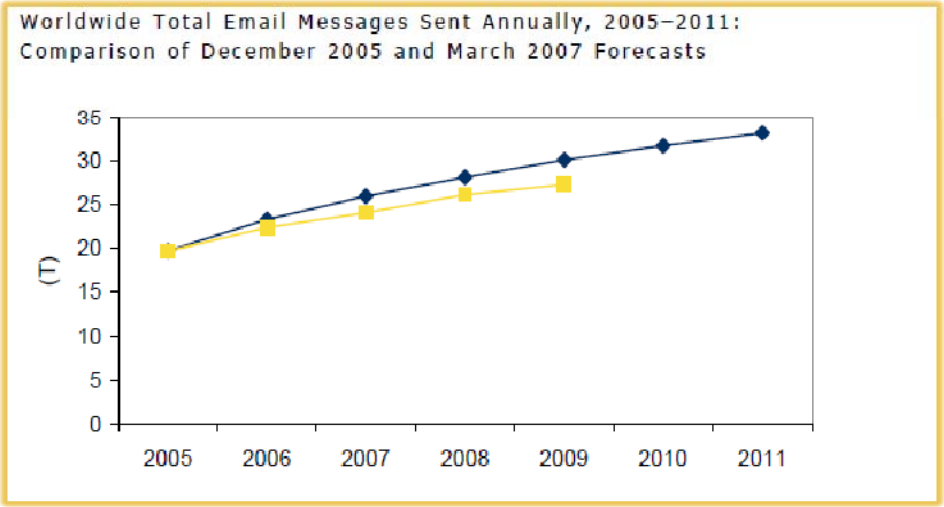
Building Your Big Data Future with Open Source

Openwave Messaging Vision

Email Growth

- By 2012 the number of mailboxes will exceed 2B.
- Total messages sent and average size of messages sent is on a steep rise

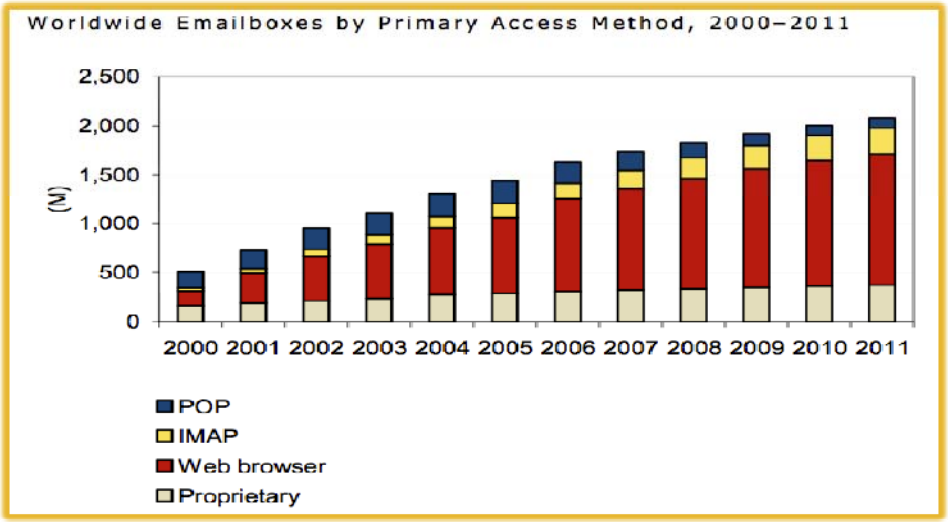
IDC, March 2010



Webmail is Dominant Mail Client

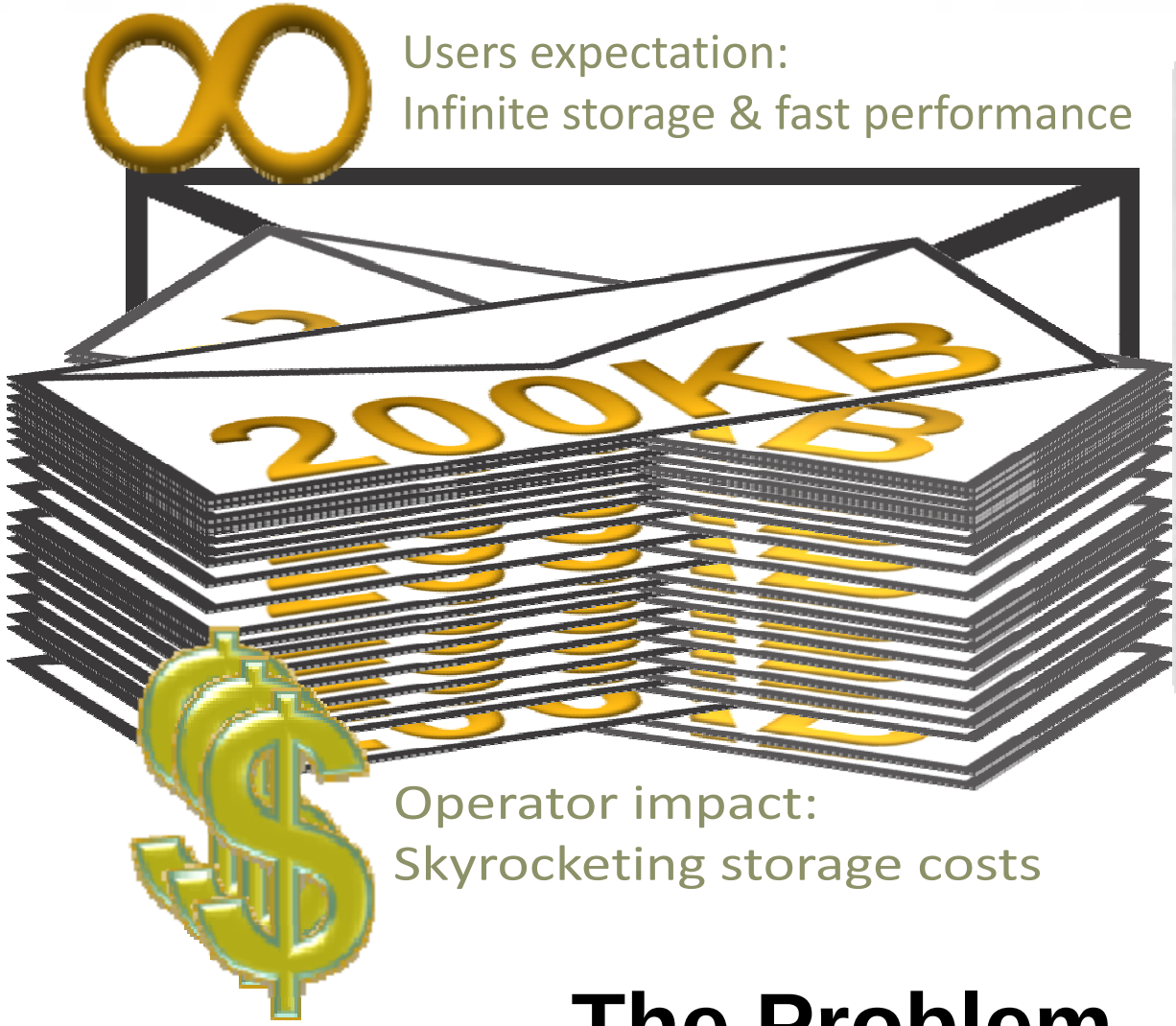
- 60% of subscribers use Webmail or IMAP as their primary access method; this will grow to 65% by 2012.

IDC, March 2010



Growth in Email Traffic and Web-based access

The Problem



Users expectation:
Infinite storage & fast performance

Operator impact:
Skyrocketing storage costs

The Problem

Consumers want

- Infinite, Searchable Inbox
- Ubiquitous access (IMAP, Webmail …)
- Highly reliable service
- And of course, FREE…

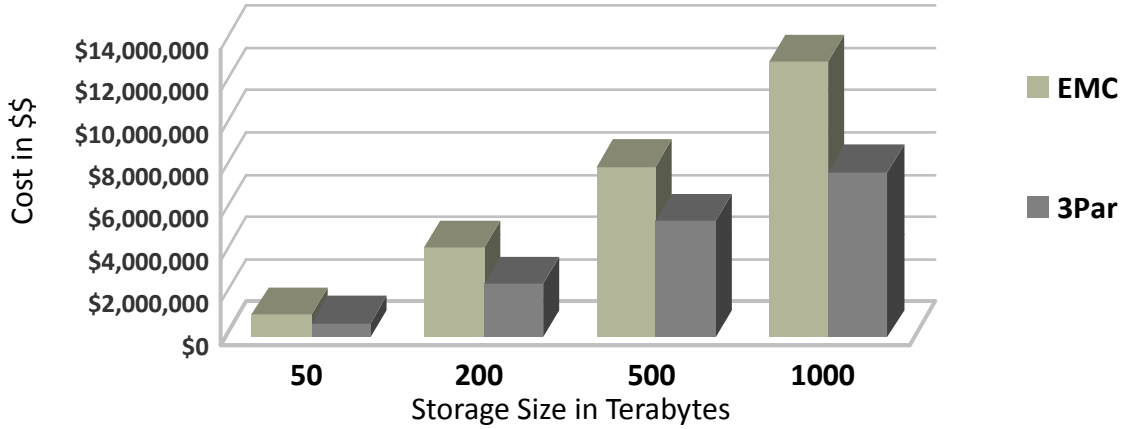
- Storage TCO becomes a huge issue

Service Providers need

- Low-TCO solution that allows providers to offer:
  - "Infinite", searchable message store without going bankrupt
  - Provides differentiation in terms of **service quality/reliability** and user experience
- Geo-Redundancy / Disaster Recovery
  - Service continuity is becoming more important than ever

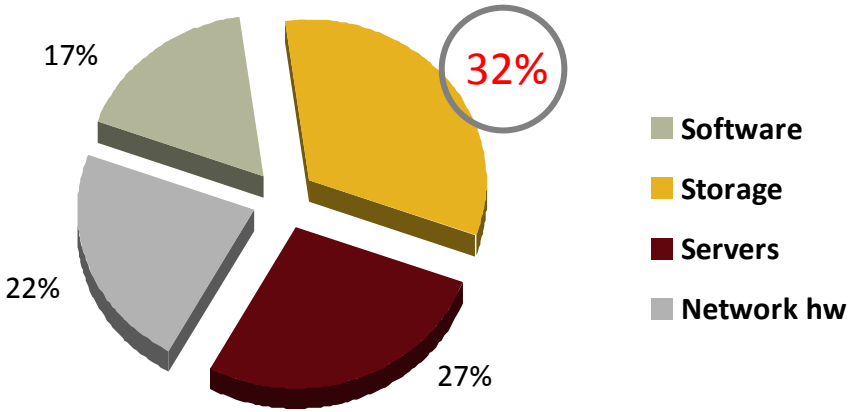
Building Your Big Data Future with Open Source

COST OF STORAGE SPACE



Increase in average mailbox size leads to very high storage needs

COSTS DISTRIBUTION FOR OPERATOR



Typical operator spends 32% of their budget on storage

Building Your Big Data Future with Open Source

Business Objectives

- **TCO reduction**
  - >50% reduction in CAPEX
- **Five-9s Reliability**
  - Active/Active Geo-Redundancy for Diaster Recovery
- **Make it scale to peta-bytes**
- **Performance goals**
  - >50% improvement over previous releases

How we achieved them

- **Picked Cassandra**
  - License free, highly scalable
- **Re-designed for Active / Active Geo-Redundancy**
  - Cassandra supported that out of the box
- And most of all, reduce **IOPS**

Why Cassandra?

- Open source, Commercially supported
- Virtually unlimited scalability
- Runs on commodity hardware
- Highly fault tolerant
- Built-in support for Geo-Redundancy
- No single point of failure – no centralized control node
- Very high read/write performance
- Self Healing (for the most part) ☺

Lessons learnt

- **Unlearn SQL, Learn NoSQL**
  - Hard to give up relational DB design
- **Use Cassandra's native faculties**
  - Don't expect ACID, its not Oracle ☺
- **Test, tune, repeat…**

•

Questions?