

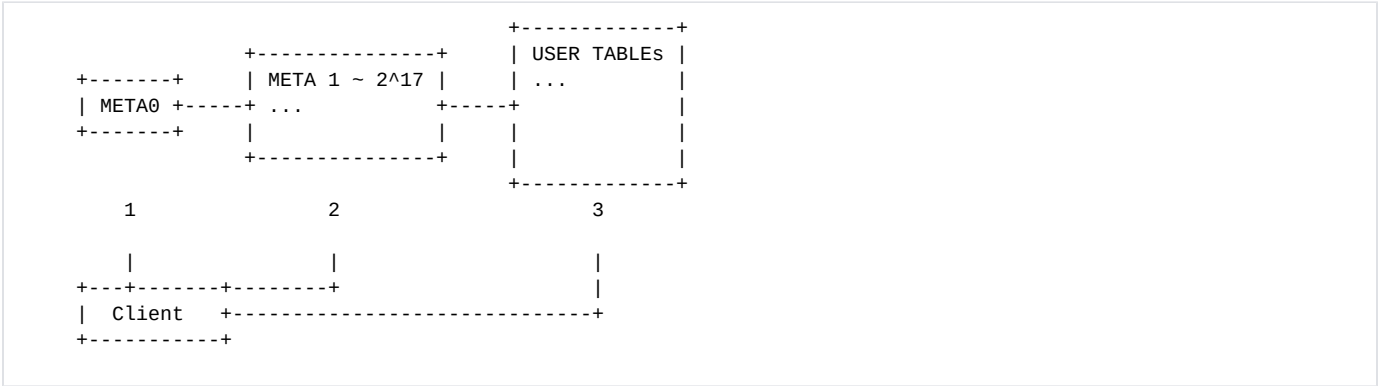
# HbaseForWebMail

## Goals

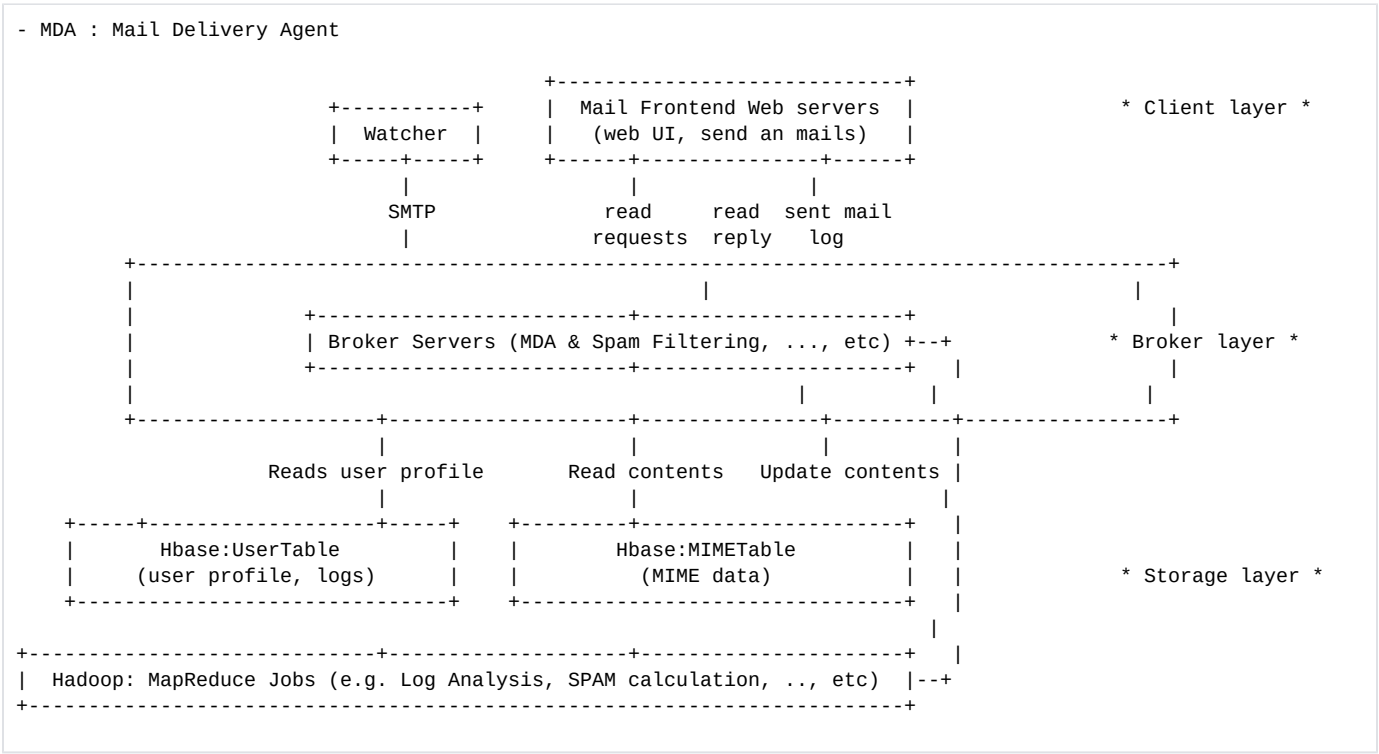
System should be :

- Stable & Reliable
- Fault tolerant
- Scalability
- Efficient
- Cost-effective
- Maintainable

## Table Indices



## Block Diagram



## Table Schema

Row Key is a email address, a unique identifier and a Sender & Receiver. Column families are meta data. (e.g. from, to, cc, bcc, reply-to, ..., etc)

```

=====
From:
To:
=====
edwardyoon@apache.org      From:edward@udanax.org <"Edward J. Yoon">
                             To:chanwit@gmail.com <"Chanwit">
edward@udanax.org          From:hama-dev@incubator.apache.org <"hama-dev">

Subject:                    Content:                    Address-Book:                    ...
-----
<"Hello, I'm ed">          <"This is a content">    Address-Book:doug@apache.org <"Doug Cutting">    ...
                             Address-Book:channy@gmail.com <"Channy Yoon">
                             ...

```

## SPAM filtering

- Sender Reputation
- Social Graph
- Text Pattern