ADITI LEAD
Learning and Development

presents

Open Talk Series

A series of illuminating talks and interactions that open our minds to new ideas and concepts; that makes us look for newer or better ways of doing what we did; or point us to exciting things we have never done before. A range of topics on Technology, Business, Fun and Life.

Be part of the learning experience at Aditi.

Join the talks. Its free.
Free as in freedom at work, not free-beer.

Its not training. Its mind-opener.

Speak at these events. Or bring an expert/friend to talk.

Mail OpenTalk@aditi.com with topic and availability.

HOW TO ENJOY AN OPEN TALK

Bring coffee & friends

Switch OFF mobile

Switch ON mind

Sign attendance sheet

SHARE your wisdom

QUESTION notions

THANK the Talker

SPREAD the good word

Sundararajan Subramanian

Image Copyright : facebook

facebook in 20 Minutes

Statistics

What is Facebook

Technical challenges

Front End

Data arch

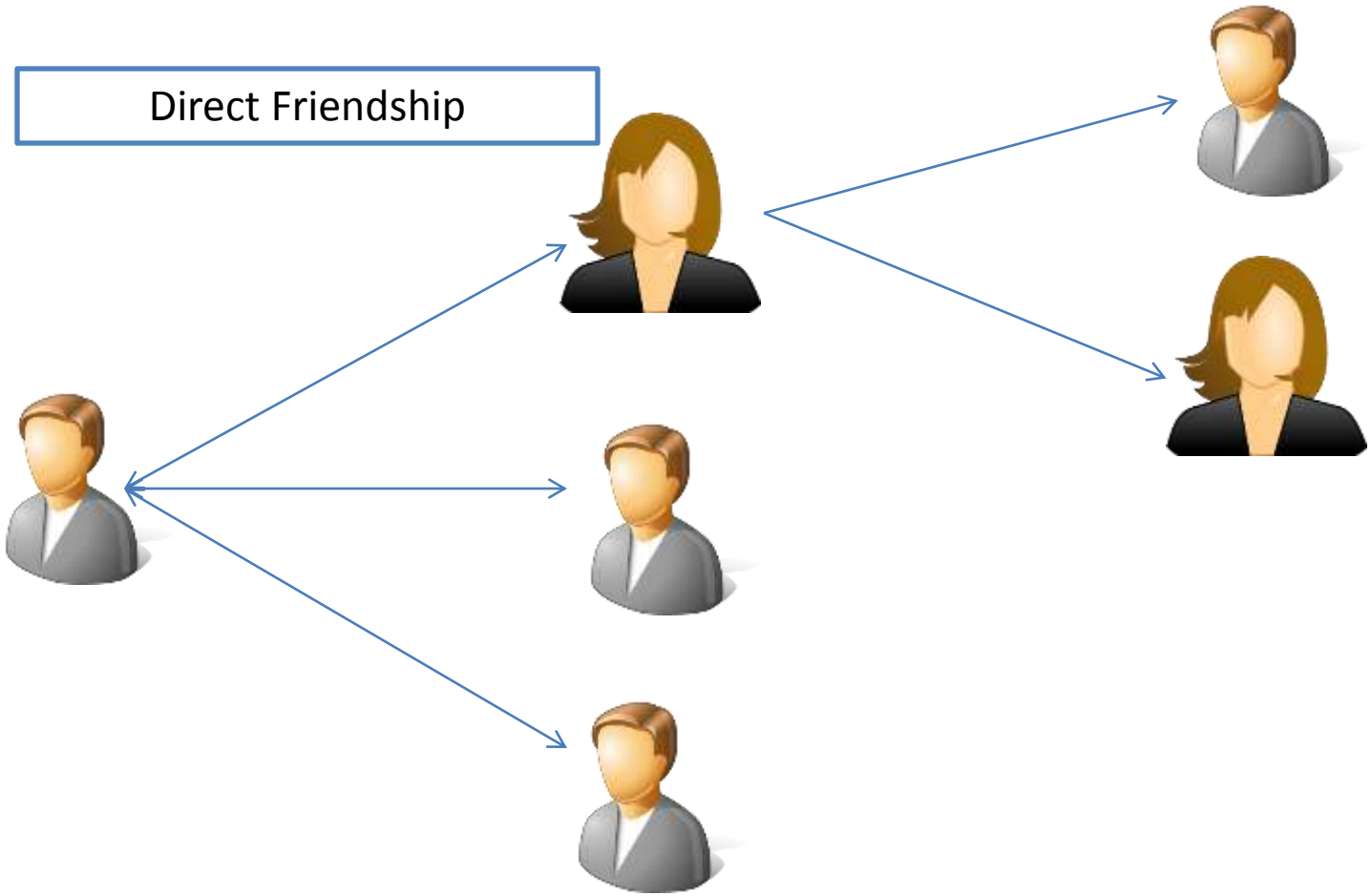
Services architecture

- 2.7 M Photos
- 10.2M Comments
- 4.6 Messages

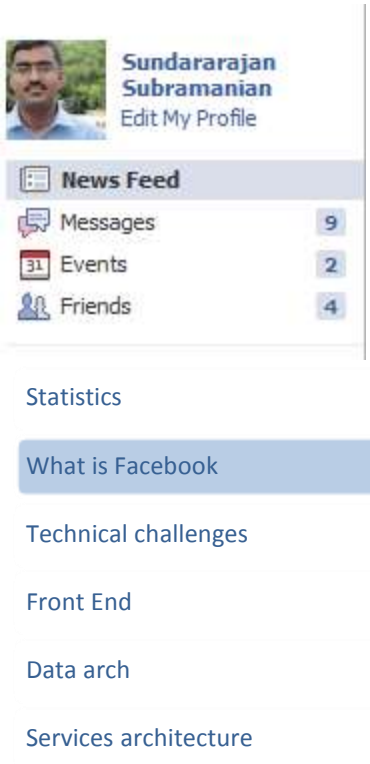
- Shared links: 1,000,000
- Tagged photos: 1,323,000
- Event invites sent out: 1,484,000
- Wall Posts: 1,587,000
- Status updates: 1,851,000
- Friend requests accepted: 1,972,000
- Photos uploaded: 2,716,000
- Comments: 10,208,000
- Message: 4,632,000

facebook in 20 Minutes

- Statistics
- What is Facebook
- Technical challenges
- Front End
- Data arch
- Services architecture



What is facebook



- A social graph
- Friends , Friends of friends, somewhere in the network.
- Friends can comment, like, read your posts
- Friends of friends can just read
- Facebook messages – chat/ email/ SMS
- Near real-time updates

Technical Challenges

Challenges

- High Concurrency
- High Data Volumes
- Multilevel Hierarchical data

Ok to Live with

- Not Mission Critical
- Cached data is fine
- Write Failures are tolerable

The Data – (Illustrational)

Everything is a hash lookup

User ID	Friends with	User Name	Age	Bio	Interests
1	2,3,4	XYZ	..	..	..
2	1	..	..	..	..

Challenges

- The Relational Nature of the data
- Data Volumes
- Concurrency

Solutions

- No Constraints, No Joins in MySQL
- Write Through cache implementation
- Hash Ring based architecture

facebook – Data Partition initial thoughts

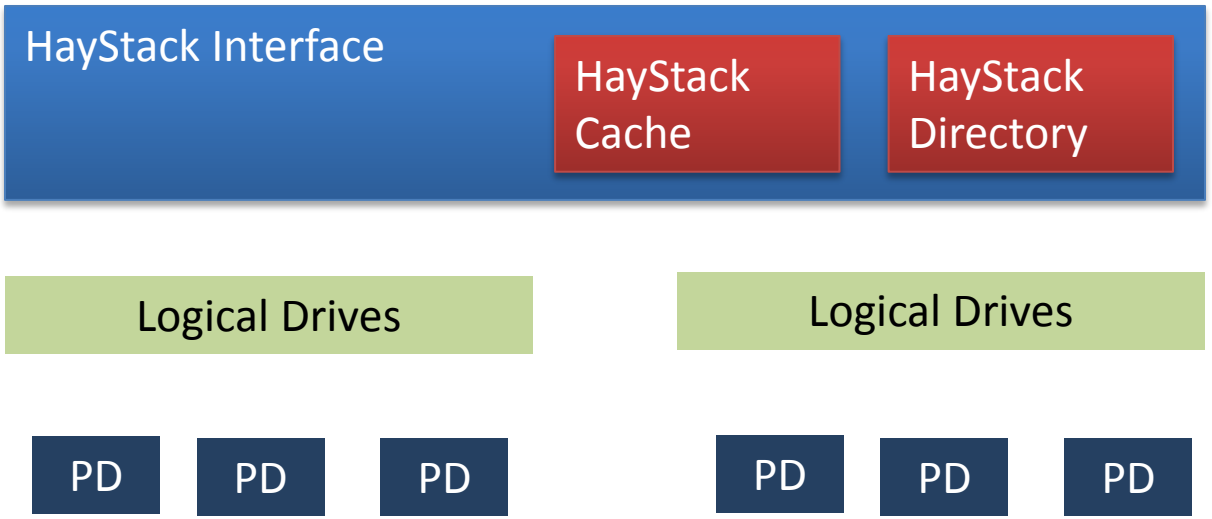
Statistics
What is Facebook
Technical challenges
Front End
Data arch
Services architecture

- Horizontal partitioning based on Networks.
  - Harvard
  - Stanford
  - Carnegie

facebook –Photos - HayStack

- Each File read required a minimum of 3 i/o in a typical file system
- CDNs- Not a Solution
- Haystack is a customized storage system, which minimizes the amount of File metadata and involves only 1 i/o for each file read.
- Haystack caches extensive data in in its main memory

facebook – HayStack



http://CDN/Cache/Machine id/(Logical volume, Photo)

Facebook – Serving the Photo - Haystack

Statistics

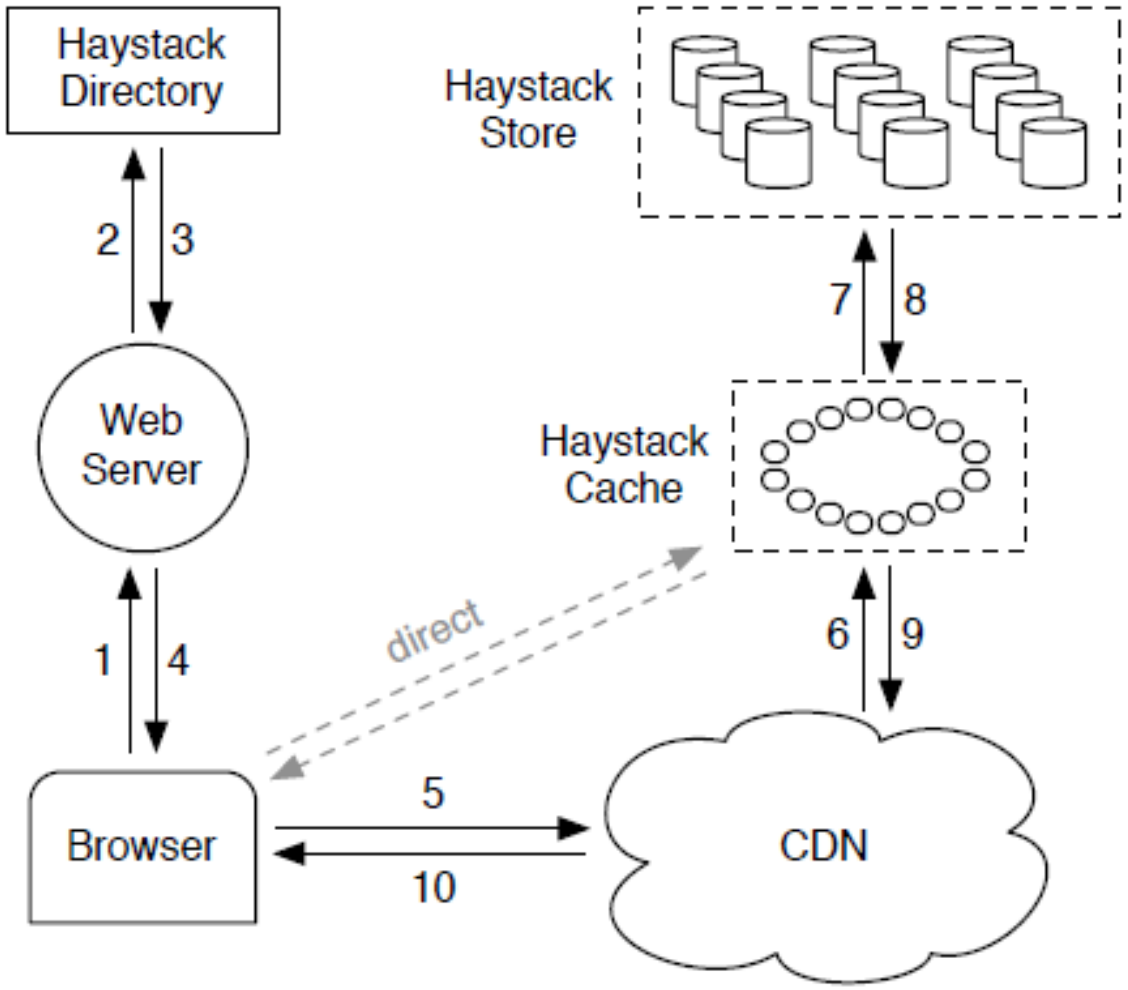
What is Facebook

Technical challenges

Front End

Data arch

Services architecture

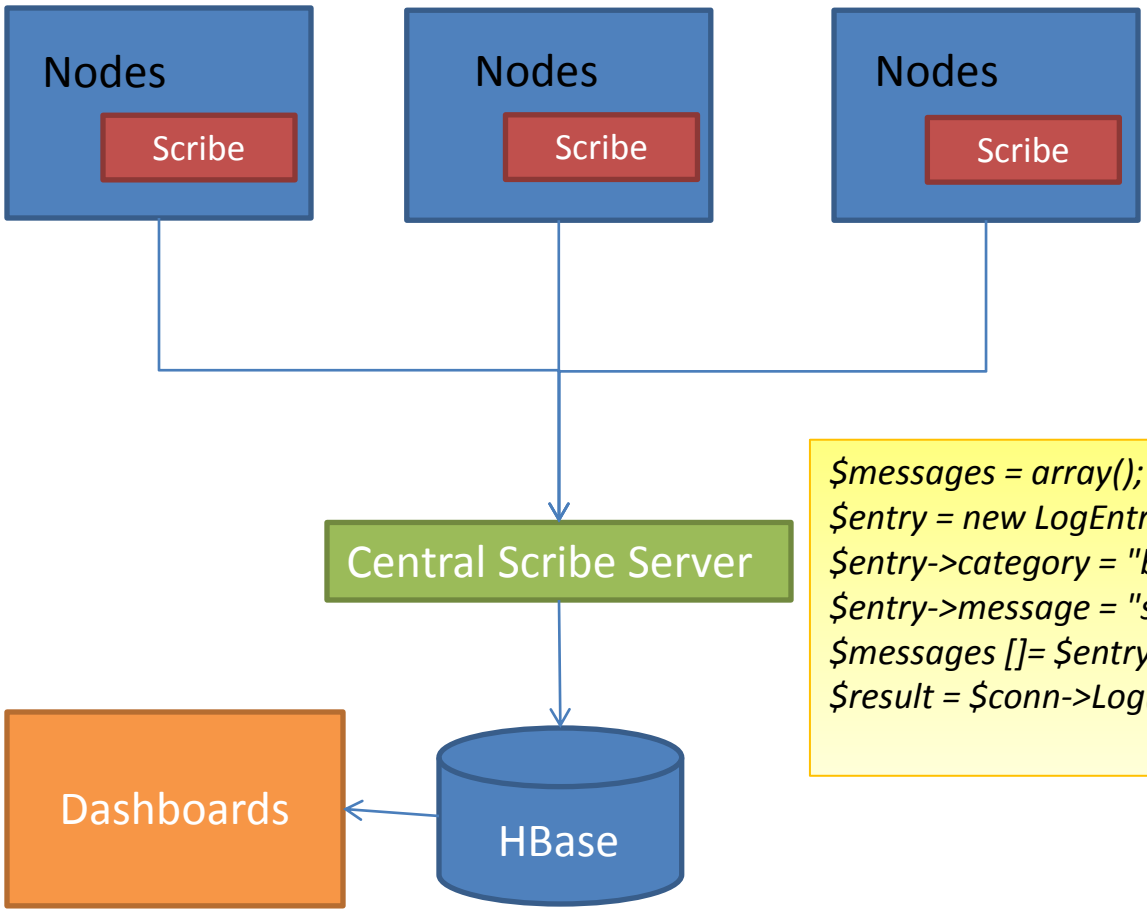


Facebook – Scribe - Logging

Sundararajan Subramanian
Edit My Profile

- News Feed
- Messages 9
- Events 2
- Friends 4

- Statistics
- What is Facebook
- Technical challenges
- Front End
- Data arch
- Services architecture




```
$messages = array();
$entry = new LogEntry;
$entry->category = "buckettest";
$entry->message = "something very";
$messages []= $entry;
$result = $conn->Log($messages);
```

facebook – Services– Thrift

- Lightweight Software framework for cross-language development
- Dev need not worry about serialization , connection handling and threading
- Supported bindings:
  - C++, PHP, Python, java, ruby, erlang, perl, haskell
- Transports : Simple interface to i/o
- Protocols : Serialization format
  - TBinaryProtocol, TJsonProtocol
- Severs
  - Non Blocking, Async, Single threaded, multi-threaded

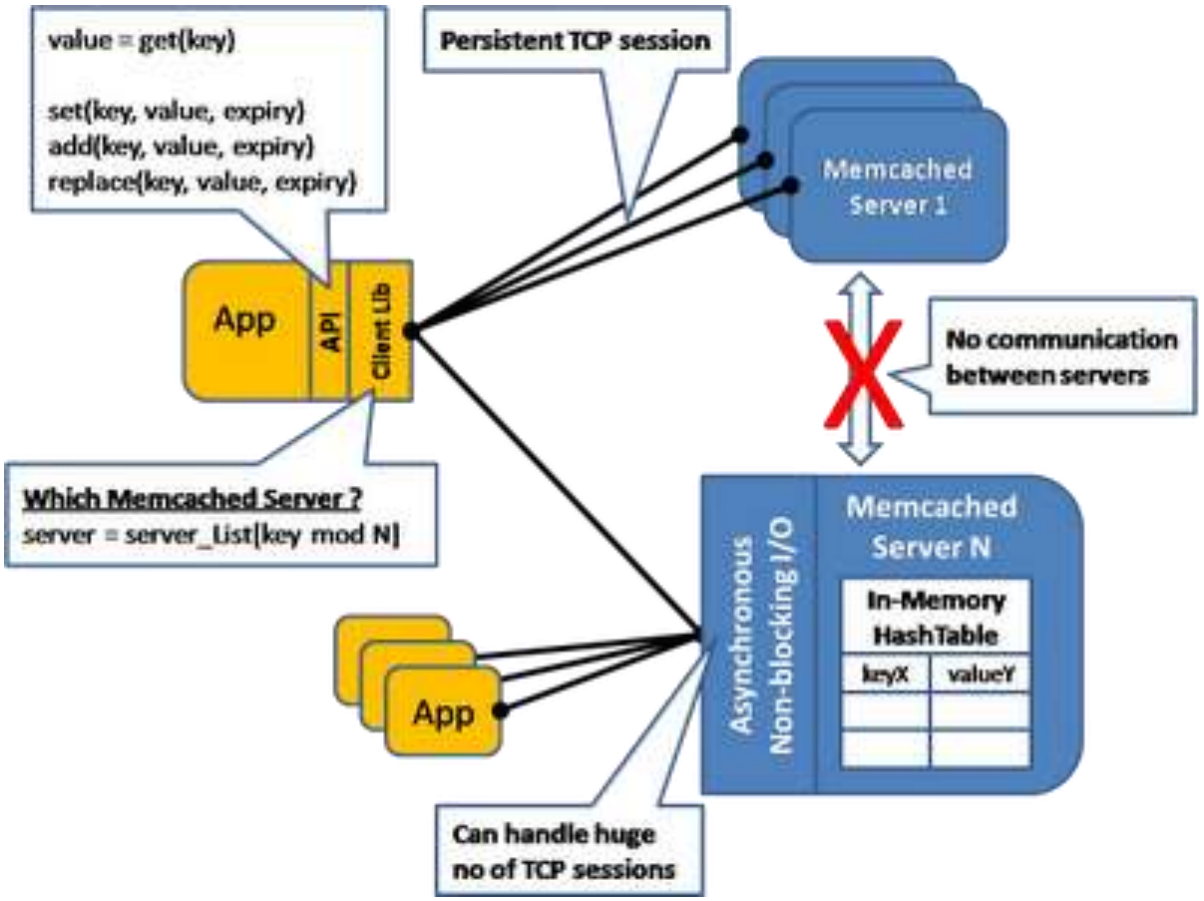
facebook – Memcache

Sundararajan Subramanian
Edit My Profile

News Feed
Messages 9
Events 2
Friends 4

Statistics
What is Facebook
Technical challenges
Front End
Data arch
Services architecture

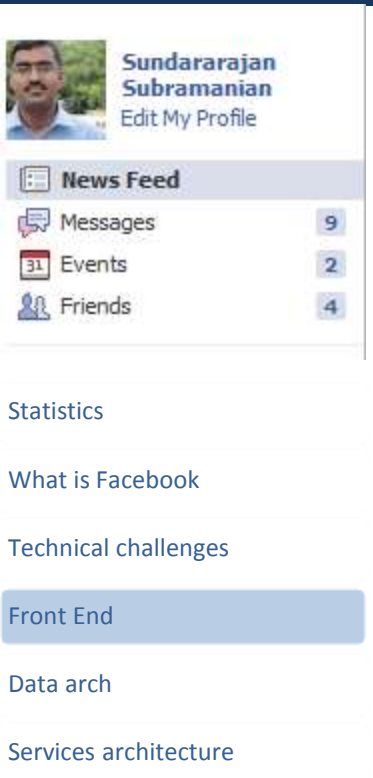
- In-memory distributed hash table
- "hot" data from MySQL stored in cache



facebook – front end - PHP

- Op – Code Optimization
- APC improvements(alternate PHP cache)
  - Lazy Loading
  - Cache priming
- Custom Extensions
  - Memcache Client Extension
  - Serialization format
  - Logging, Stats Collection, Monitoring
  - Asynchronous event-handling mechanism

facebook – front end – Hip Hop



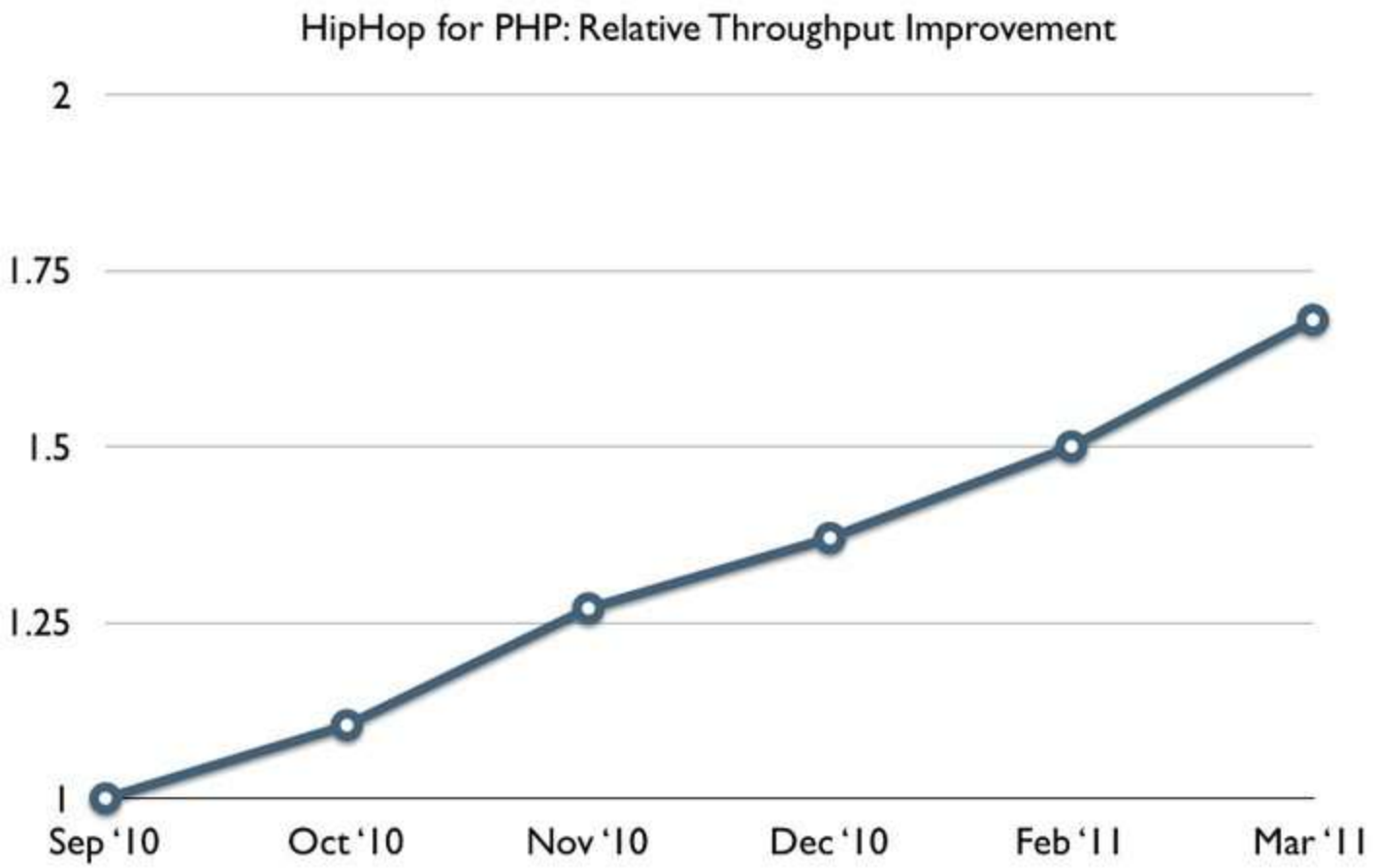
- Source Code Transformer
- Static Analysis, type inference, Code Generation
- Easier to write extensions
- Significantly cuts down on CPU and Memory usage

facebook – front end – Hip Hop

Sundararajan Subramanian
Edit My Profile

- News Feed
- Messages 9
- Events 2
- Friends 4

- Statistics
- What is Facebook
- Technical challenges
- Front End
- Data arch
- Services architecture



facebook – front end – BigPipe

Sundararajan Subramanian
Edit My Profile

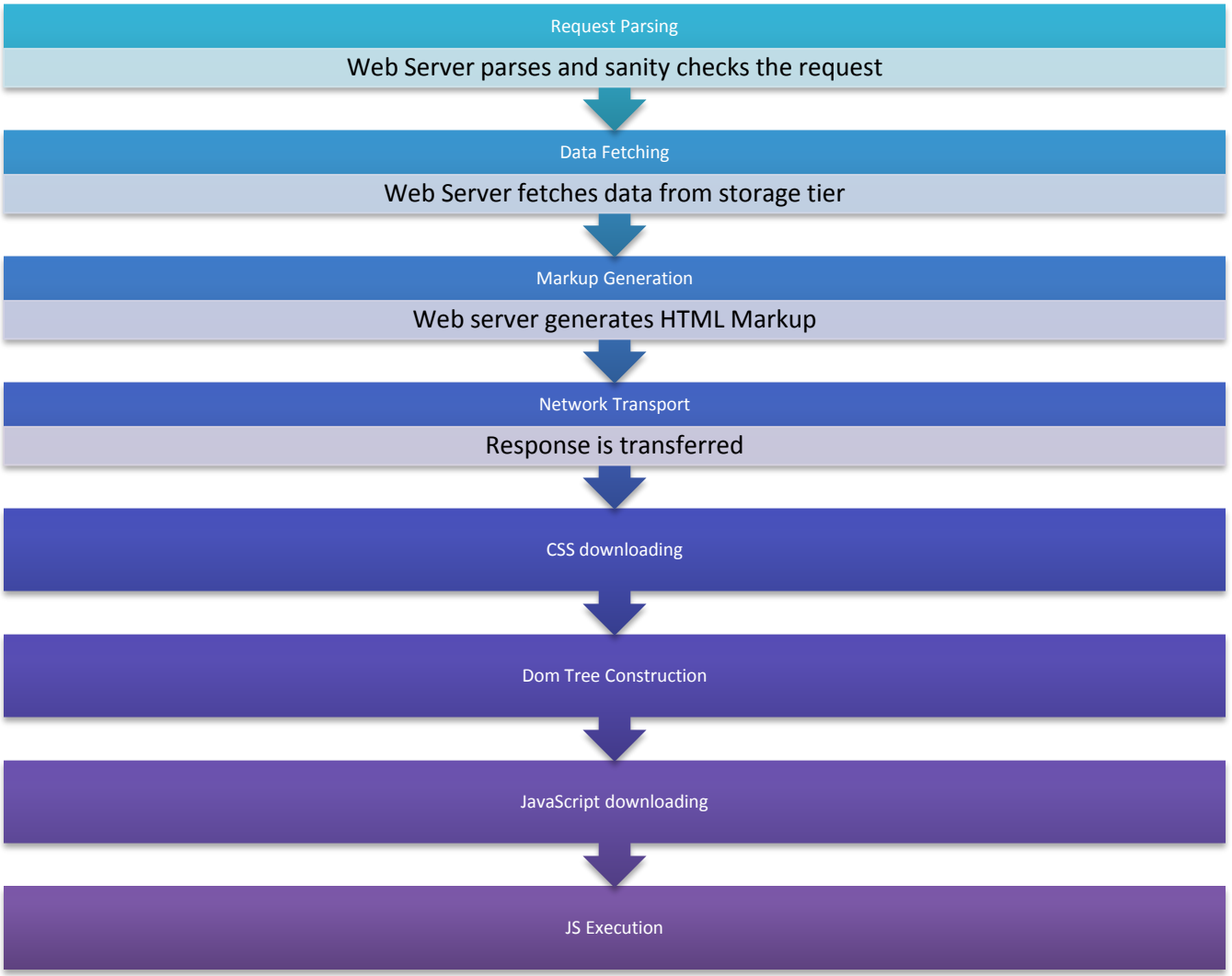
- News Feed
- Messages 9
- Events 2
- Friends 4

- Statistics
- What is Facebook
- Technical challenges
- Front End
- Data arch
- Services architecture

BigPipe first breaks web pages into multiple chunks called pagelets

facebook – front end – BigPipe

BigPipe first breaks web pages into multiple chunks called pagelets



facebook – Technology Stack

Front End

Big Pipe

Hip Hop

PHP - Custom compiler / Cache implementations

Linux – Custom Kernel Extensions

Scribe

Thrift

Service Aggregators

Service 1

Service 2

Service 3

Service 4

Data Store

MemCache – Write Through Cache implementation

Cassandra

MySQL

HBase

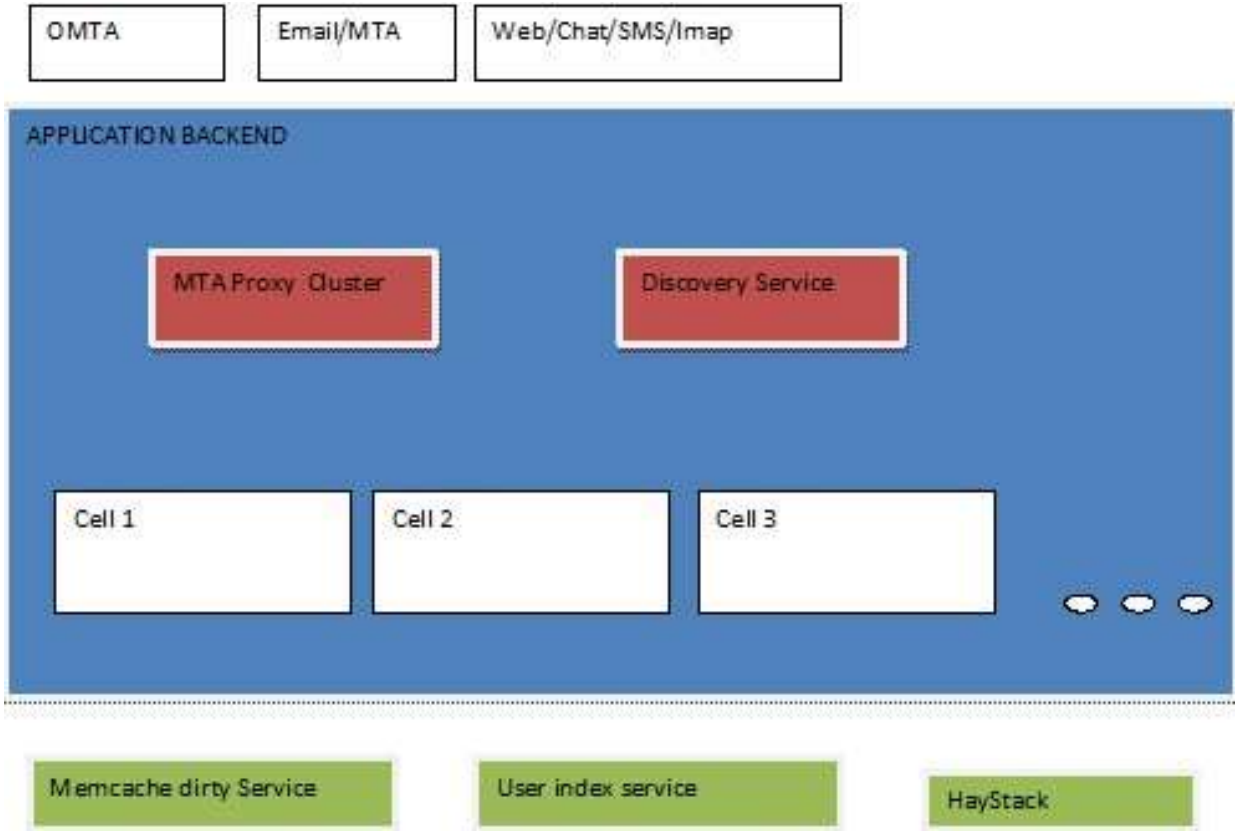
HayStack

facebook – Messages Infrastructure

Sundararajan Subramanian
Edit My Profile

- News Feed
- Messages 9
- Events 2
- Friends 4

- Statistics
- What is Facebook
- Technical challenges
- Front End
- Data arch
- Services architecture
- Messages



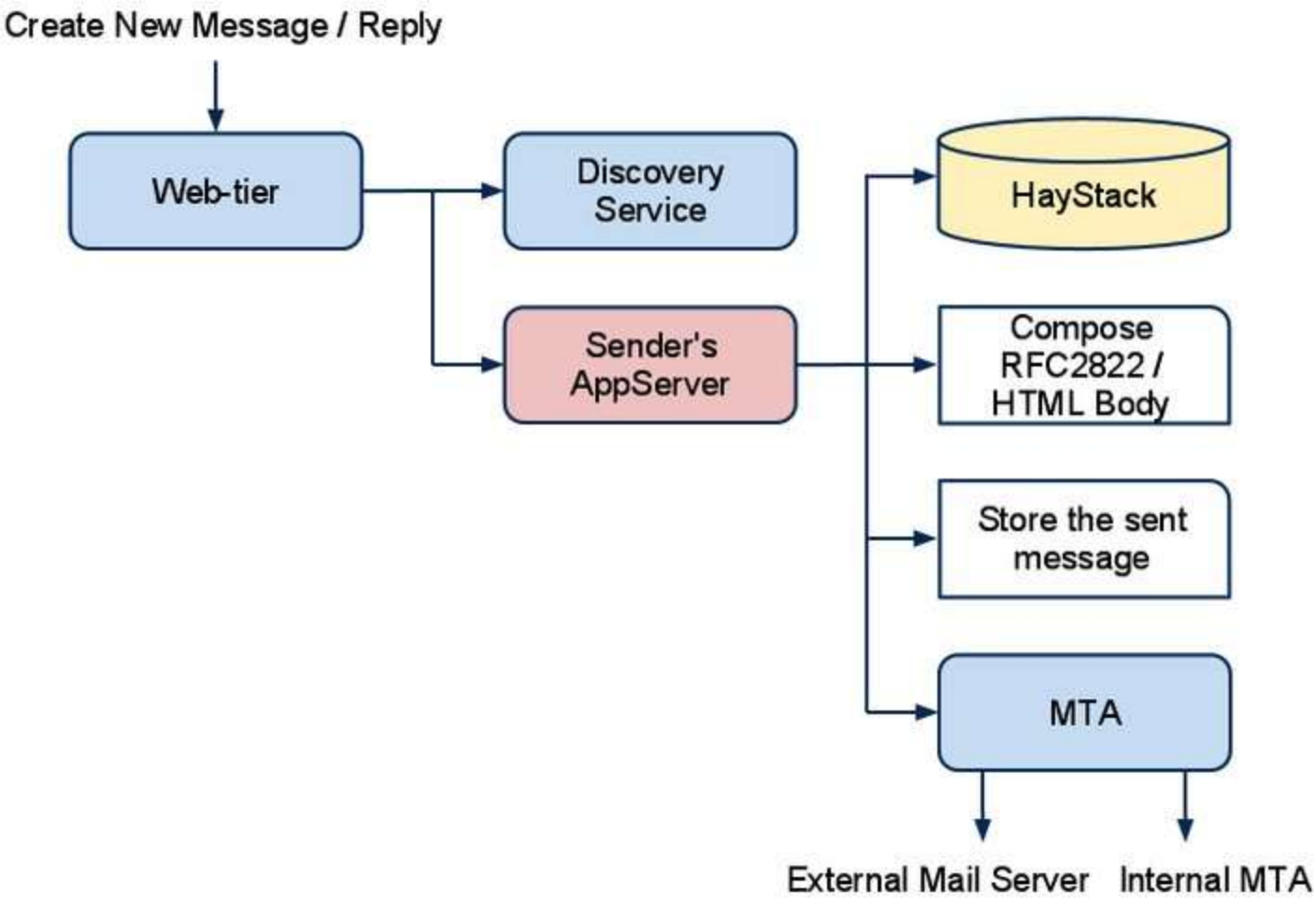
facebook - Messages

Sundararajan Subramanian
Edit My Profile

- News Feed
- Messages 9
- Events 2
- Friends 4

- Statistics
- What is Facebook
- Technical challenges
- Front End
- Data arch
- Services architecture
- Messages

Outgoing Flow

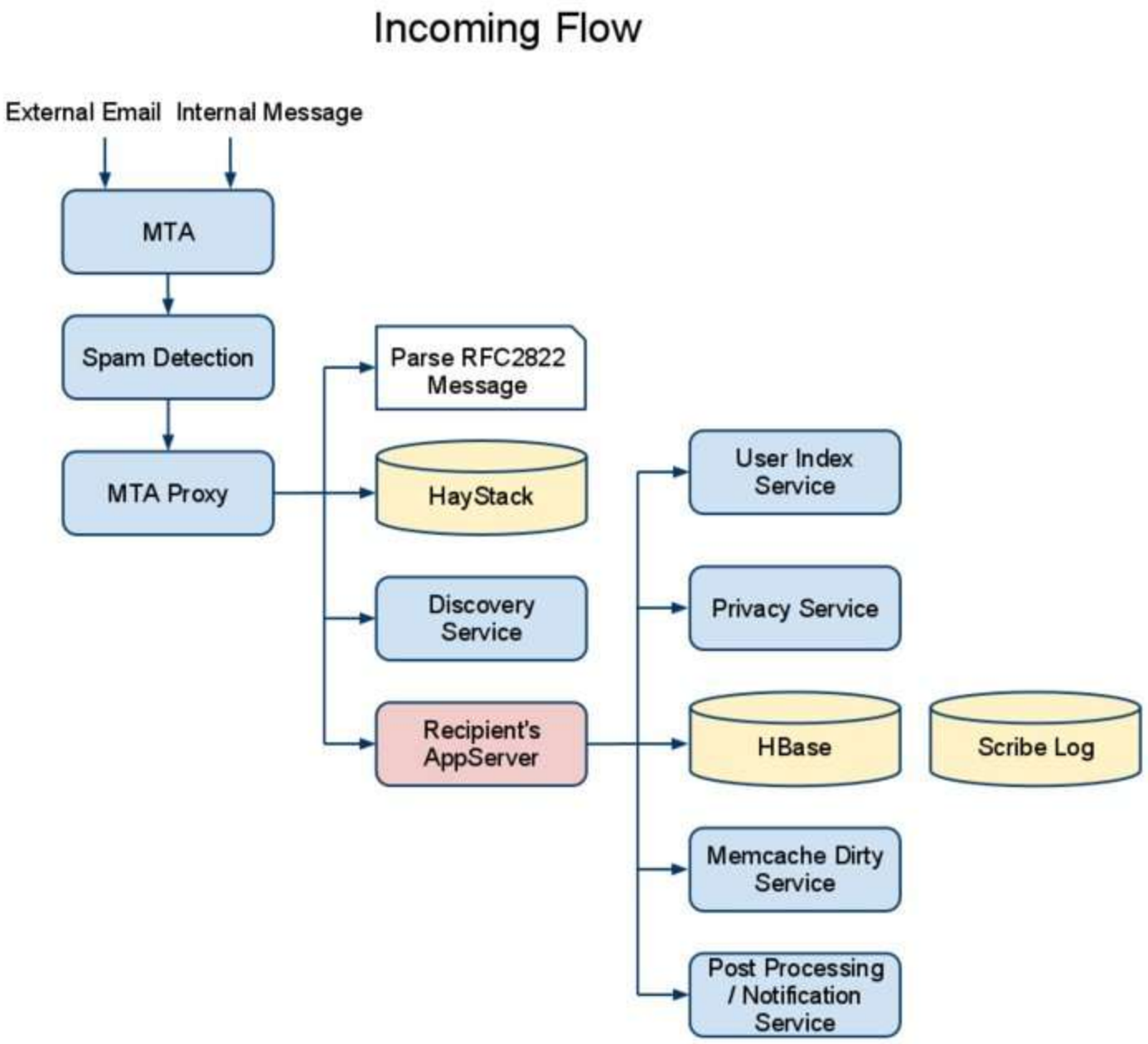


facebook - Messages

Sundararajan Subramanian
Edit My Profile

- News Feed
- Messages 9
- Events 2
- Friends 4

- Statistics
- What is Facebook
- Technical challenges
- Front End
- Data arch
- Services architecture
- Messages



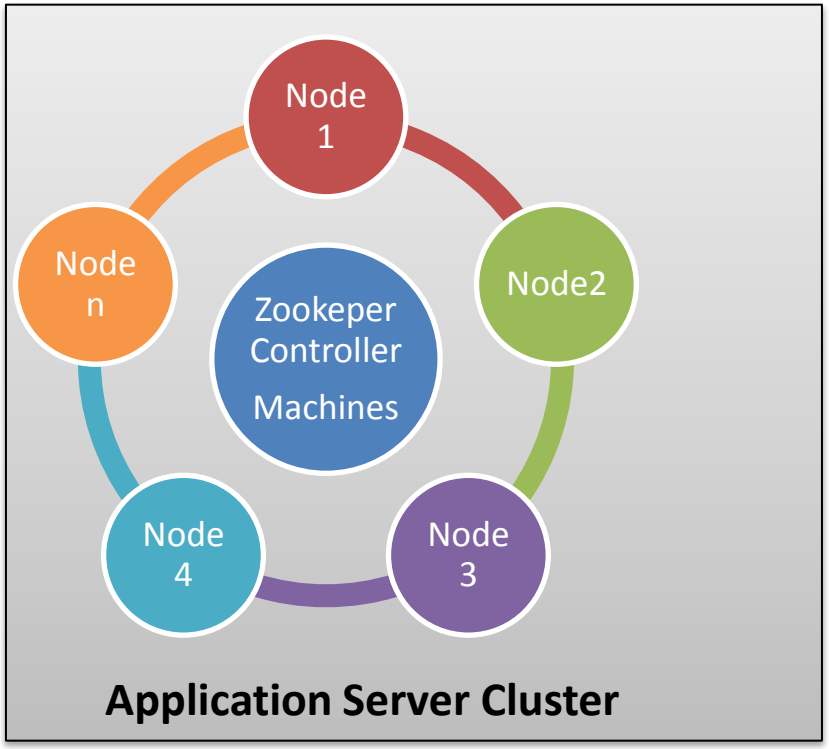
facebook – Cells

Sundararajan Subramanian
Edit My Profile

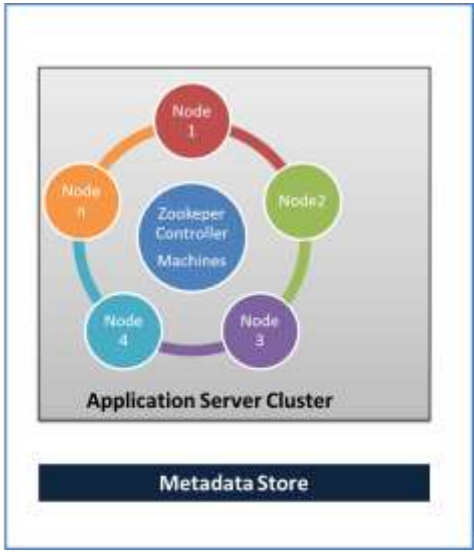
- News Feed
- Messages 9
- Events 2
- Friends 4

- Statistics
- What is Facebook
- Technical challenges
- Front End
- Data arch
- Services architecture
- Messages

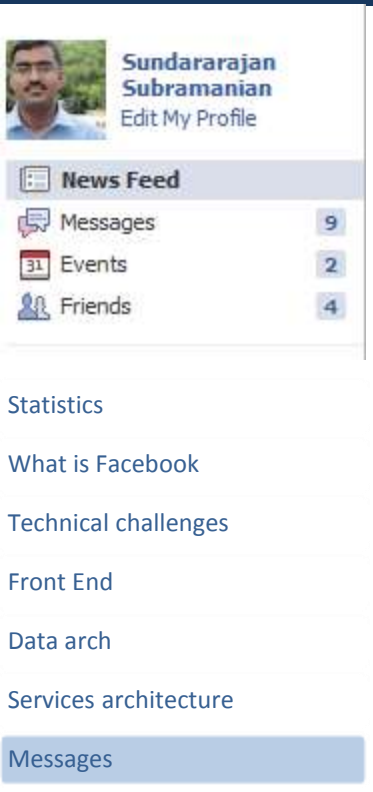
Cell



Metadata Store



facebook – Cells



- They help scale incrementally while limiting failure scenarios
- Easy upgrades
- Metadata store failures affect only a few users
- Easy rollout
- Flexibility to host cells in different data centers with multi-homing for disaster recovery

Take away – for our applications

- Really parallel Asynchronous AJAX Pages
  - ASP.Net Update panels is a HOAX
- Appropriate usage of client side technology
- Cache – Cache – Cache
  - Write Through Caches are way better
  - App Fabric cache/ Memcache
- High – Normalization is not needed
  - Store denormalized views – materialized views
- Parallel Services and Service aggregators
- Fault tolerant applications
- Asynchronous Processing
- 1 Sec response time is too SLOW

References

- http://facebook.com/engineering
- www.infoq.com
- www.highscalability.com
- www.stackoverflow.com
- www.thrift.org

LEAD Looking for speakers for
OPEN TALK
OPEN
YOU!
Sahil Sagar
ADITI
GREAT PLACE TO WORK INSTITUTE INDIA : BEST workplaces 2008
Vic Parmar
ADITI
GREAT PLACE TO WORK INSTITUTE INDIA : BEST workplaces 2008
LEAD
OPEN MIC

ADITI LEAD

Keep Learning

For suggestions on topics/ feedbacks etc.,

Contact OpenTalk@aditi.com