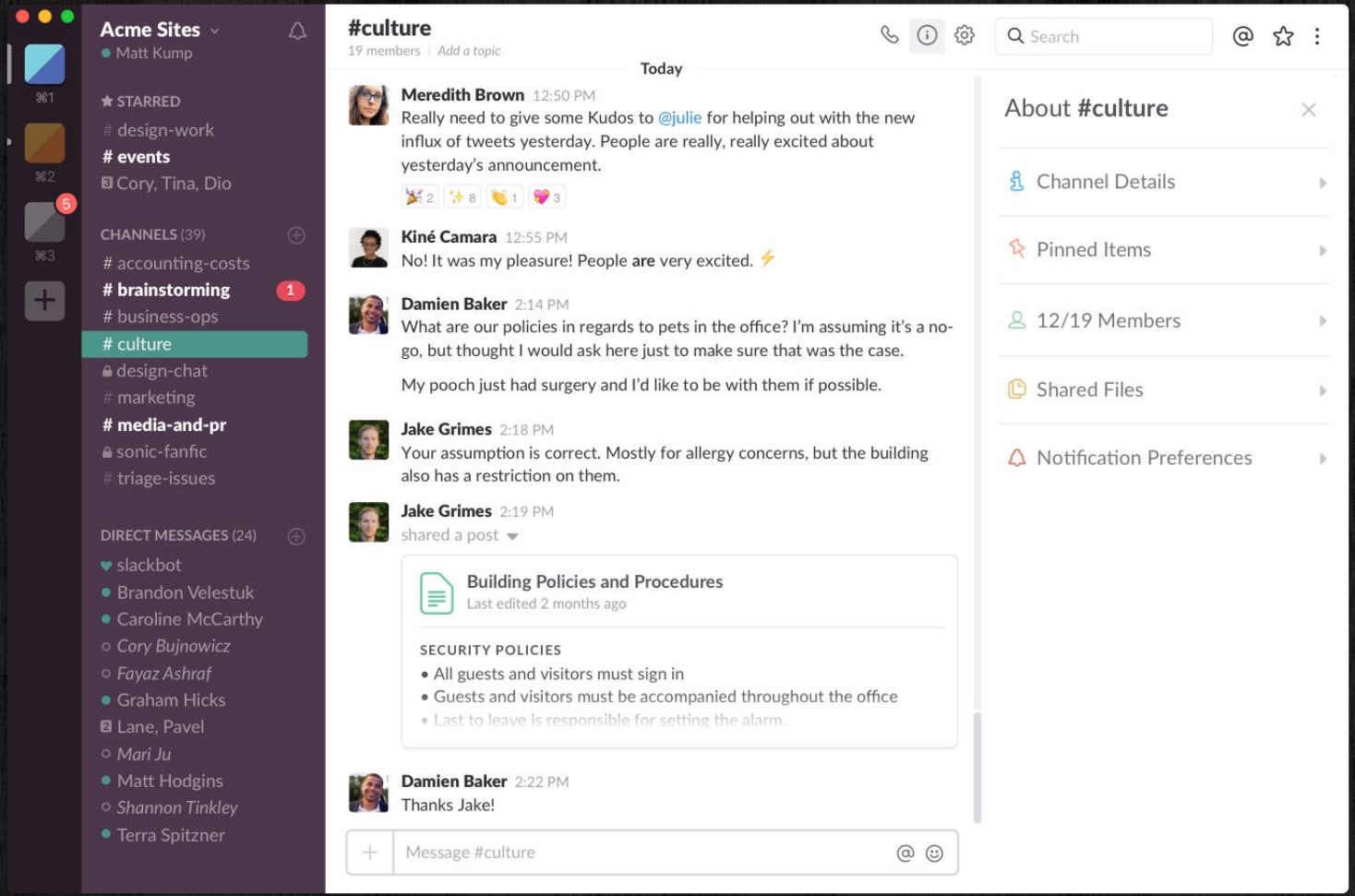
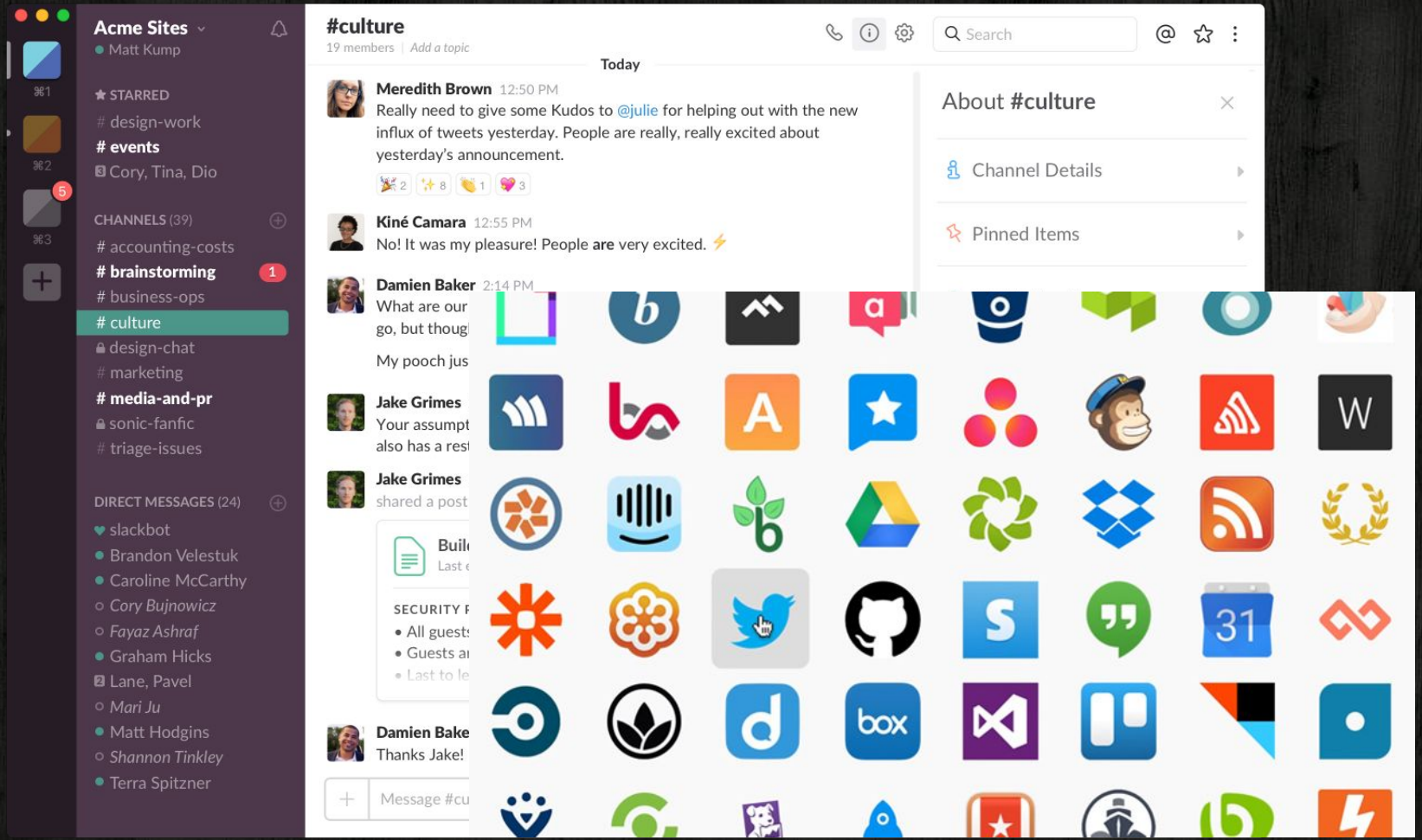
Scaling Slack

Bing Wei
Infrastructure@Slack


Acme Sites
Matt Kump
⌘1
⌘2
⌘3
STARRED
design-work
events
Cory, Tina, Dio
CHANNELS (39)
accounting-costs
brainstorming
business-ops
culture
design-chat
marketing
media-and-pr
sonic-fanfic
triage-issues
DIRECT MESSAGES (24)
slackbot
Brandon Velestuk
Caroline McCarthy
Cory Bujnowicz
Fayaz Ashraf
Graham Hicks
Lane, Pavel
Mari Ju
Matt Hodgins
Shannon Tinkley
Terra Spitzner
#culture
19 members | Add a topic
Search
Today
Meredith Brown 12:50 PM
Really need to give some Kudos to @julie for helping out with the new influx of tweets yesterday. People are really, really excited about yesterday's announcement.
Kiné Camara 12:55 PM
No! It was my pleasure! People are very excited.
Damien Baker 2:14 PM
What are our policies in regards to pets in the office? I'm assuming it's a no-go, but thought I would ask here just to make sure that was the case.
My pooch just had surgery and I'd like to be with them if possible.
Jake Grimes 2:18 PM
Your assumption is correct. Mostly for allergy concerns, but the building also has a restriction on them.
Jake Grimes 2:19 PM
shared a post
Building Policies and Procedures
Last edited 2 months ago
SECURITY POLICIES
• All guests and visitors must sign in
• Guests and visitors must be accompanied throughout the office
• Last to leave is responsible for setting the alarm.
Damien Baker 2:22 PM
Thanks Jake!
Message #culture
About #culture
Channel Details
Pinned Items
12/19 Members
Shared Files
Notification Preferences


Acme Sites
Matt Kump
STARRED
design-work
events
Cory, Tina, Dio
CHANNELS (39)
accounting-costs
brainstorming
business-ops
culture
design-chat
marketing
media-and-pr
sonic-fanfic
triage-issues
DIRECT MESSAGES (24)
slackbot
Brandon Velestuk
Caroline McCarthy
Cory Bujnowicz
Fayaz Ashraf
Graham Hicks
Lane, Pavel
Mari Ju
Matt Hodgins
Shannon Tinkley
Terra Spitzner
#culture
19 members | Add a topic
Search
Today
Meredith Brown 12:50 PM
Really need to give some Kudos to @julie for helping out with the new influx of tweets yesterday. People are really, really excited about yesterday's announcement.
Kiné Camara 12:55 PM
No! It was my pleasure! People are very excited.
Damien Baker 2:14 PM
Jake Grimes
Jake Grimes
shared a post
Damien Bake
Thanks Jake!
Message #cu
About #culture
Channel Details
Pinned Items

slack

Our Mission:

To make people's working lives simpler, more pleasant, and more productive.

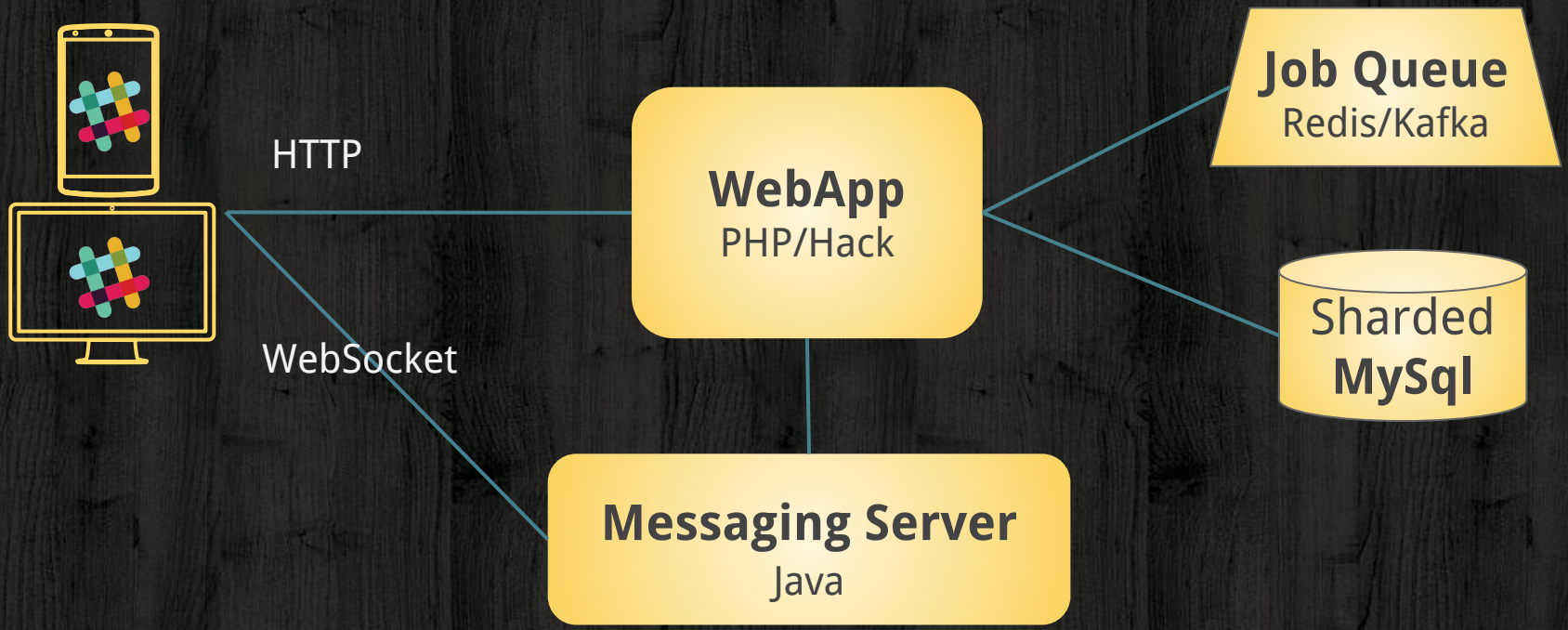
From supporting small teams

To serving gigantic organizations of hundreds of thousands of users

Slack Scale

- 6M+ DAU, 9M+ WAU
  5M+ peak simultaneously connected
- Avg 10+ hrs/weekday connected
  Avg 2+ hrs/weekday in active use
- 55% of DAU outside of US

Cartoon Architecture



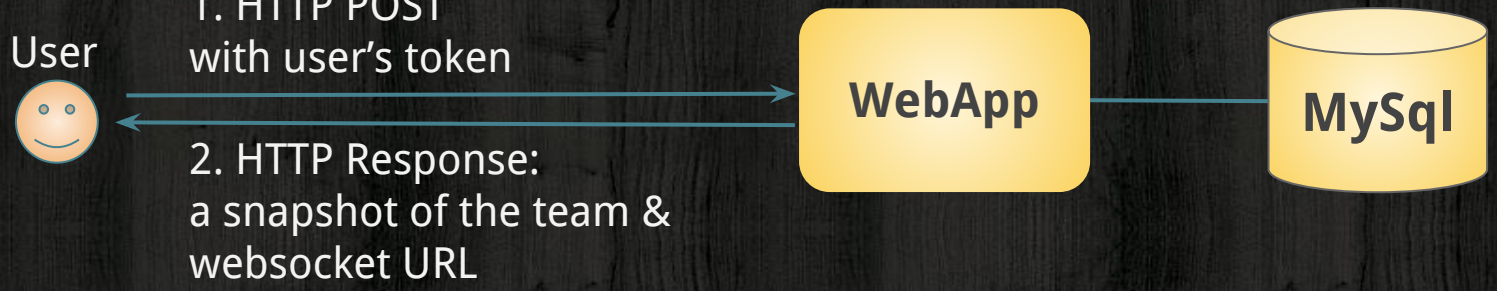
Outline

- The slowness problem
- Incremental Improvements
- Architecture changes
  - Flannel
  - Client Pub/Sub
  - Messaging Server

Challenge:

Slowness Connecting to Slack

Login Flow in 2015

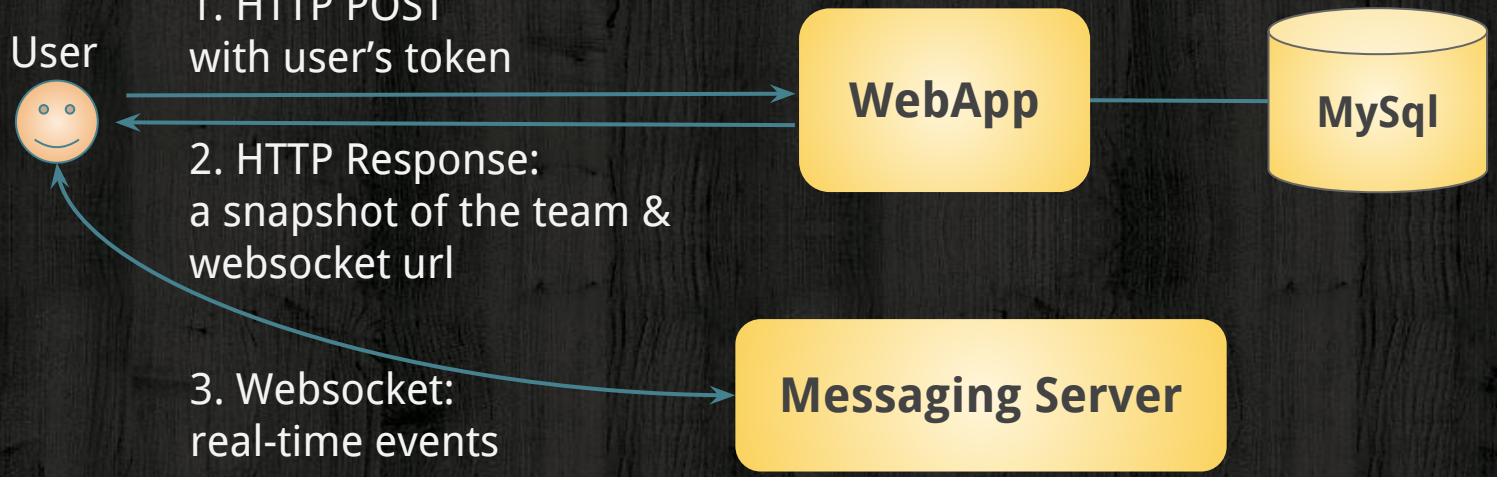


```
{
    ok: true,
  - self: {
        id: "U1WKVSERY",
        name: "binggmail",
      + prefs: {…},
        created: 1469837648,
        manual_presence: "active"
    },
  + team: {…},
    latest_event_ts: "1510364623.000000",
  + channels: […],
  + groups: […],
  + ims: […],
    cache_ts: 1510365223,
    read_only_channels: [ ],
    can_manage_shared_channels: false,
  + mpims: […],
  + subteams: {…},
  + dnd: {…},
  + users: […],
    cache_version: "v16-giraffe",
    cache_ts_version: "v2-bunny",
  + bots: […],
    url: "wss://mpmulti-qdvx.lb.slack-msgs.com/webs
    xXrCpNy2JMxpygOpoWoqLcr6WSY-alpczoEVgEYxnseZ6AA
}
```


Some examples

number of users/channels	response size
30 / 10	200K
500 / 200	2.5M
3K / 7K	20M
30K / 1K	60M

Login flow in 2015



Real-time Events on WebSocket

User

WebSocket:
100+ types of events

Messaging Server

e.g. chat messages,
typing indicator,
files uploads,
files comments,
threads replies,
user presence changes,
user profile changes,
reactions, pins, stars,
channel creations,
app installations,
etc.

Login Flow in 2015

- Clients Architecture
  - Download a snapshot of *entire* team
  - Updates trickle in through the WebSocket
  - Eventually consistent snapshot of whole team

Problems

Initial team snapshot takes time

Problems

Initial team snapshot takes time

Large client memory footprint

Problems

Initial team snapshot takes time

Large client memory footprint

Expensive reconnections

Problems

Initial team snapshot takes time

Large client memory footprint

Expensive reconnections

Reconnect storm

Outline

- The slowness problem
- Incremental Improvements
- Architecture changes
  - Flannel
  - Client Pub/Sub
  - Messaging Server

Improvements

- Smaller team snapshot
  - Client local storage + delta
  - Remove objects out + in parallel loading
  - Simplified objects: e.g. canonical Avatars

Improvements

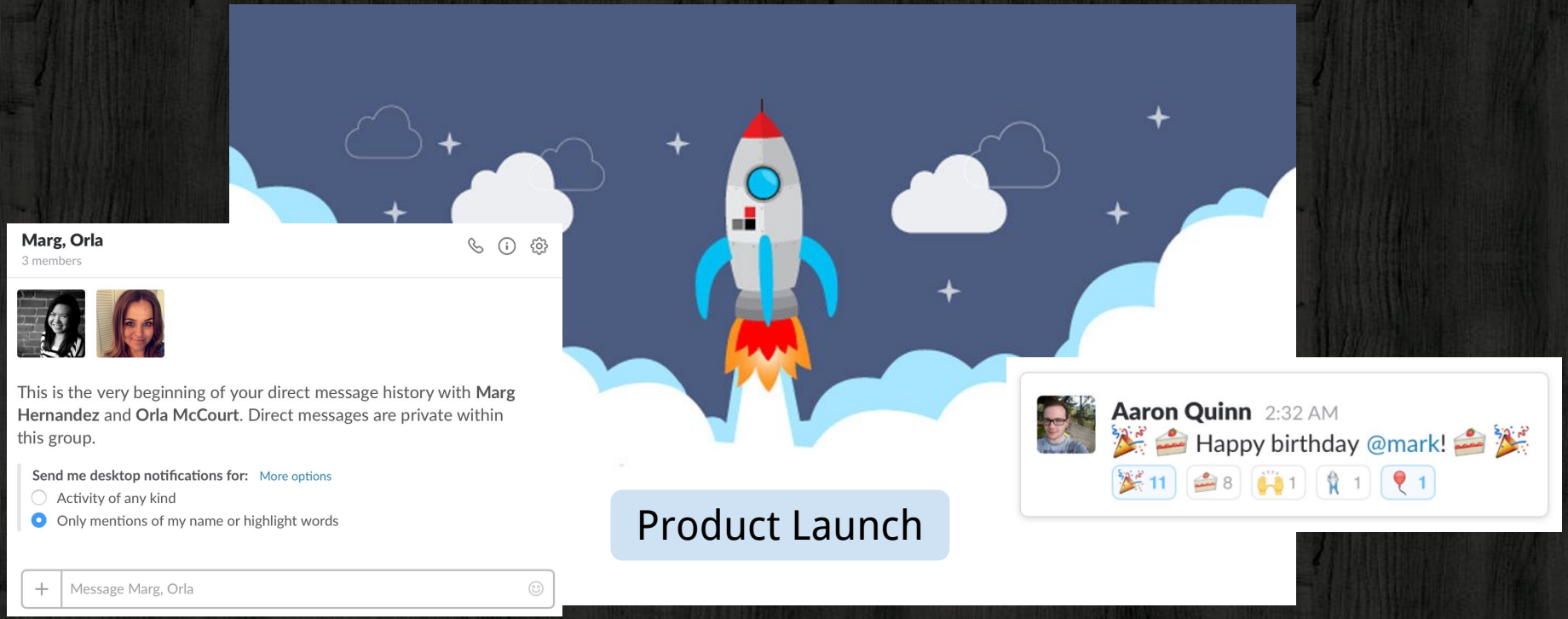
- Incremental boot
  - Load one channel first

Improvements

- Rate Limit
- POPs
- Load Testing Framework

Support New Product Features

Cope with New Product Features



Still...

Limitations

- What if team sizes keep growing
- Outages when clients dump their local storage

Outline

- The slowness problem
- Incremental Improvements
- Architecture changes
  - Flannel
  - Client Pub/Sub
  - Messaging Server

Client Lazy Loading

Download less data upfront

- Fetch more on demand

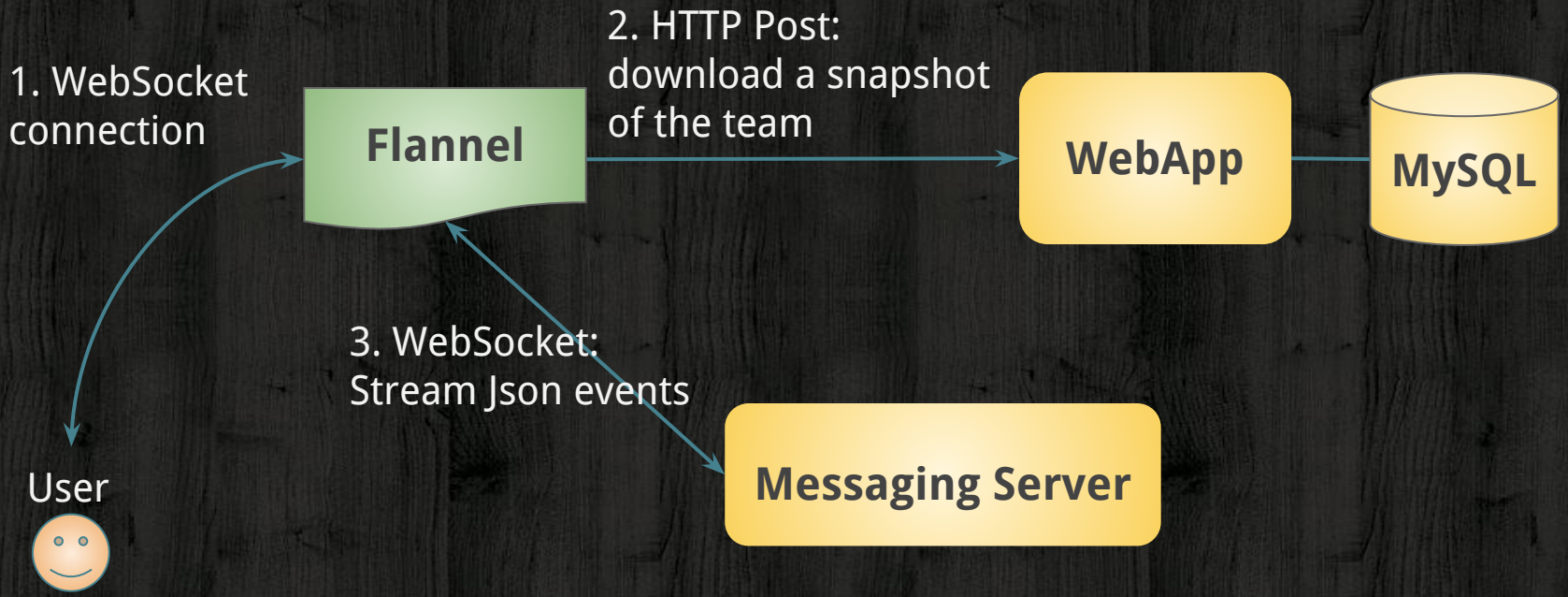
Flannel: Edge Cache Service

A query engine backed by cache on edge locations

What are in Flannel's cache

- Support big objects first
  - Users
  - Channels Membership
  - Channels

Login and Message Flow with Flannel



A Man in the Middle

Use real-time events to update its cache

E.g. user creation,

user profile change,

channel creation,

user joins a channel,

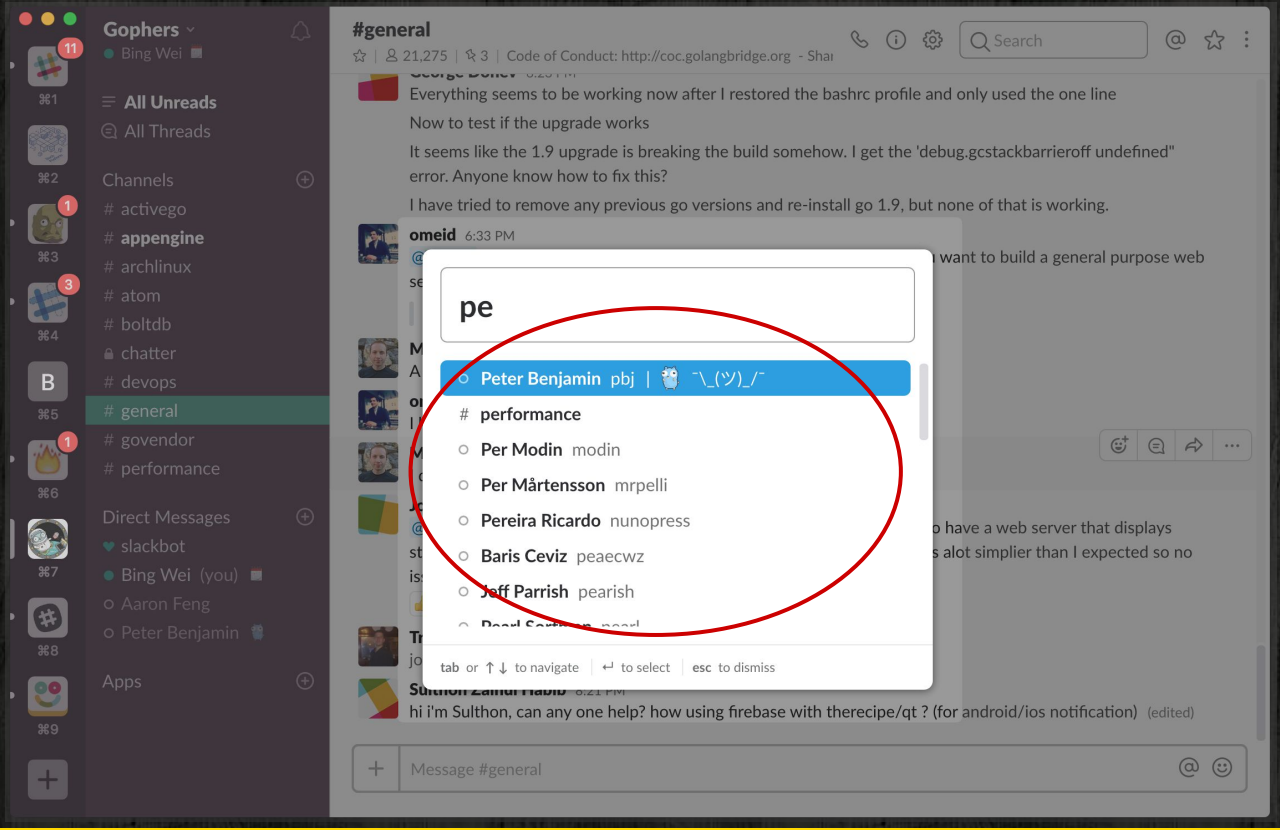
channel convert to private

Edge Locations

Mix of AWS & Google Cloud

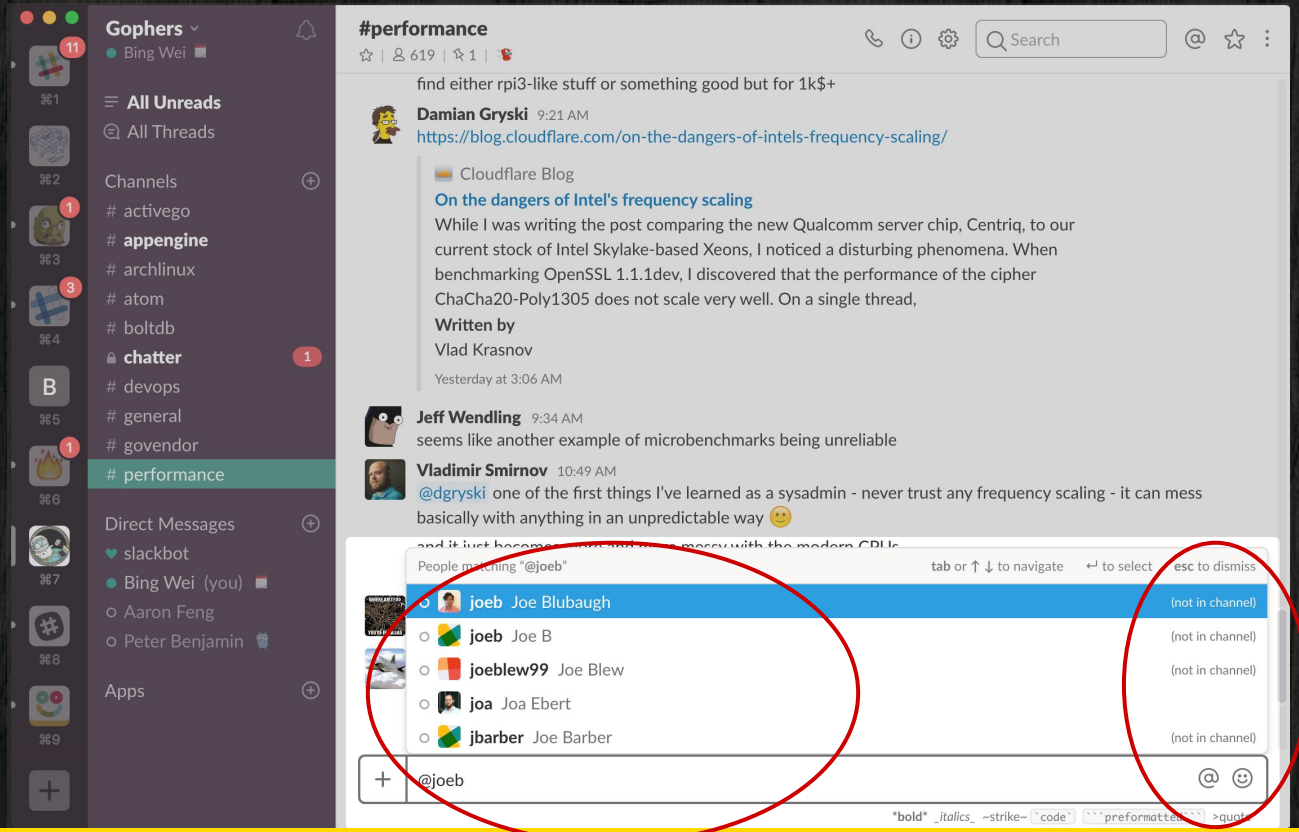
Examples Powered by Flannel

Quick Switcher



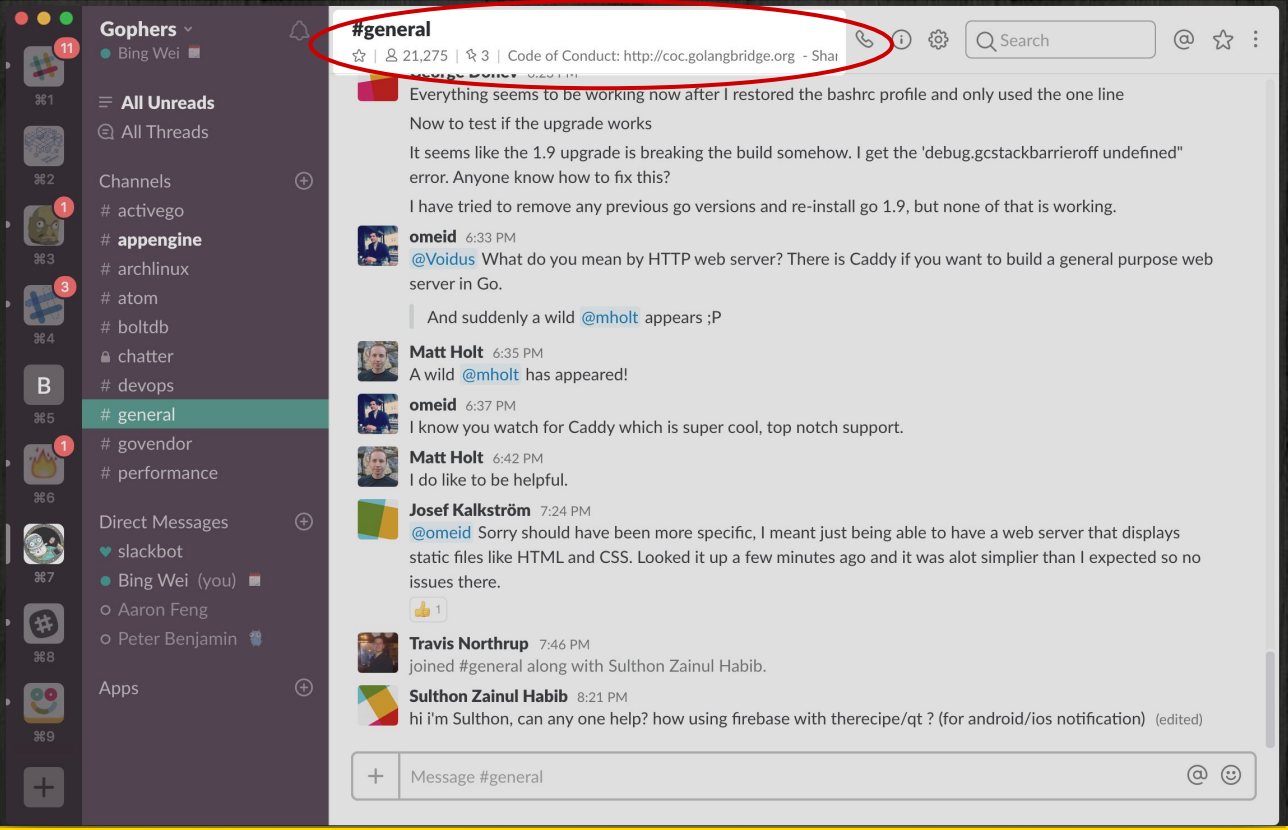
Examples Powered by Flannel

Mention Suggestion



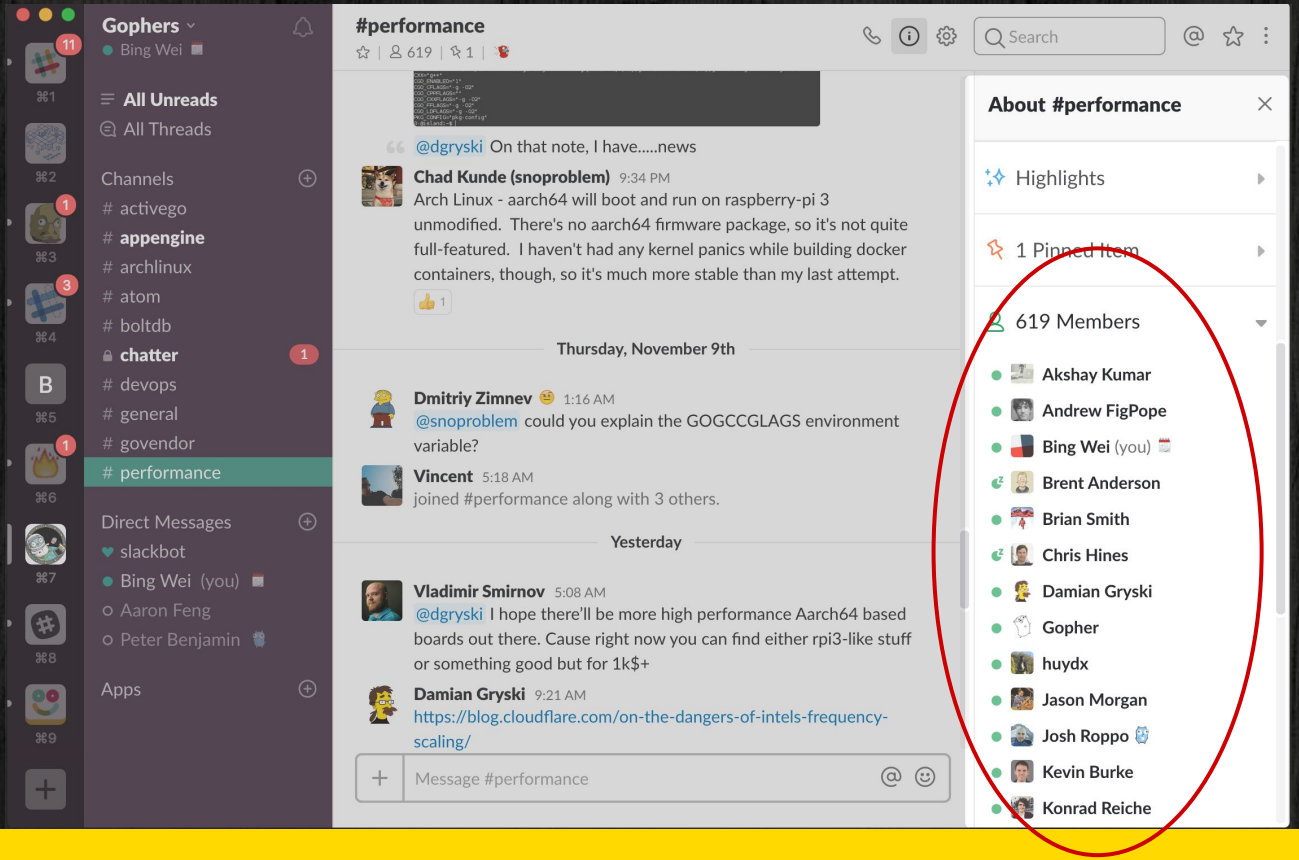
Examples Powered by Flannel

Channel Header



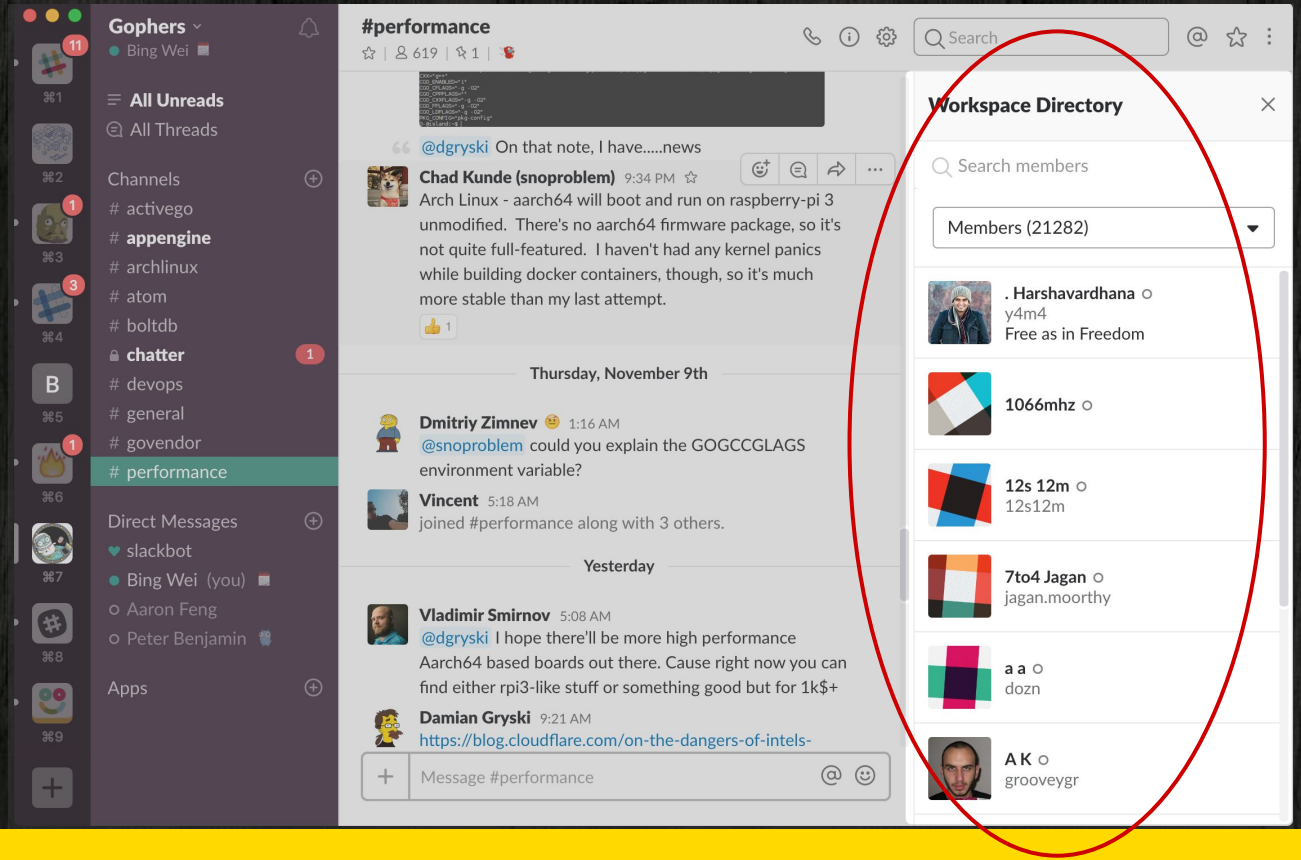
Examples Powered by Flannel

Channel Sidebar



Examples Powered by Flannel

Team Directory



Flannel Results

- Launched Jan 2017
  - Load 200K user team

- 5M+ simultaneous connections at peak
- 1M+ clients queries/sec

Flannel Results

This is not the end of the story

Evolution of Flannel

Web Client Iterations

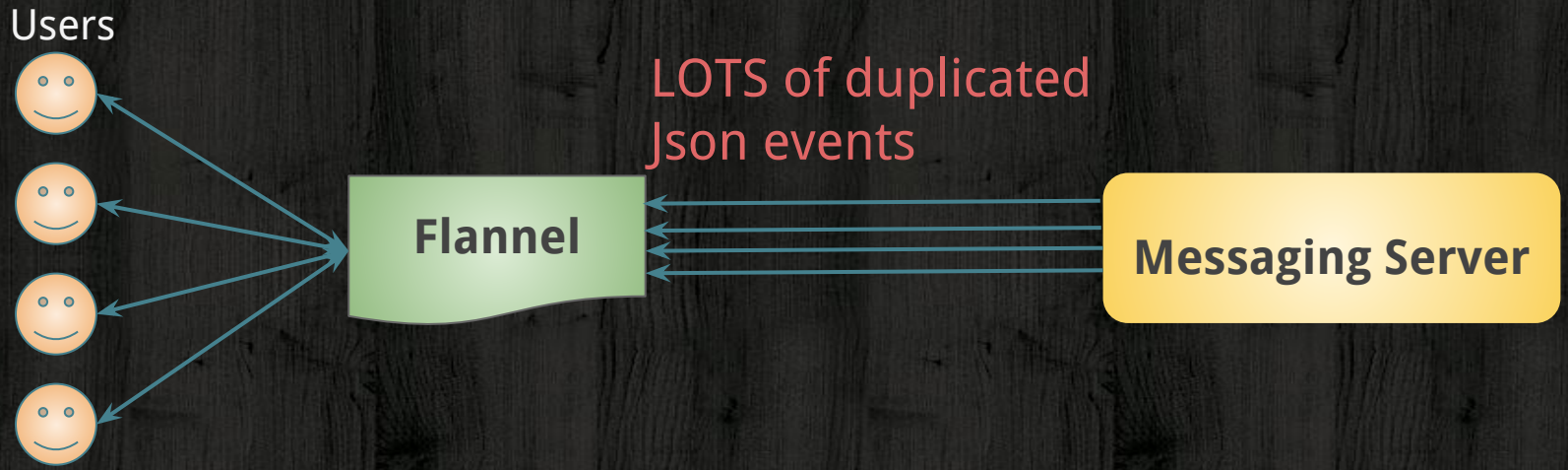
Flannel Just-In-Time Annotation

Right before Web clients are about to access an object, Flannel pushes that object to clients.

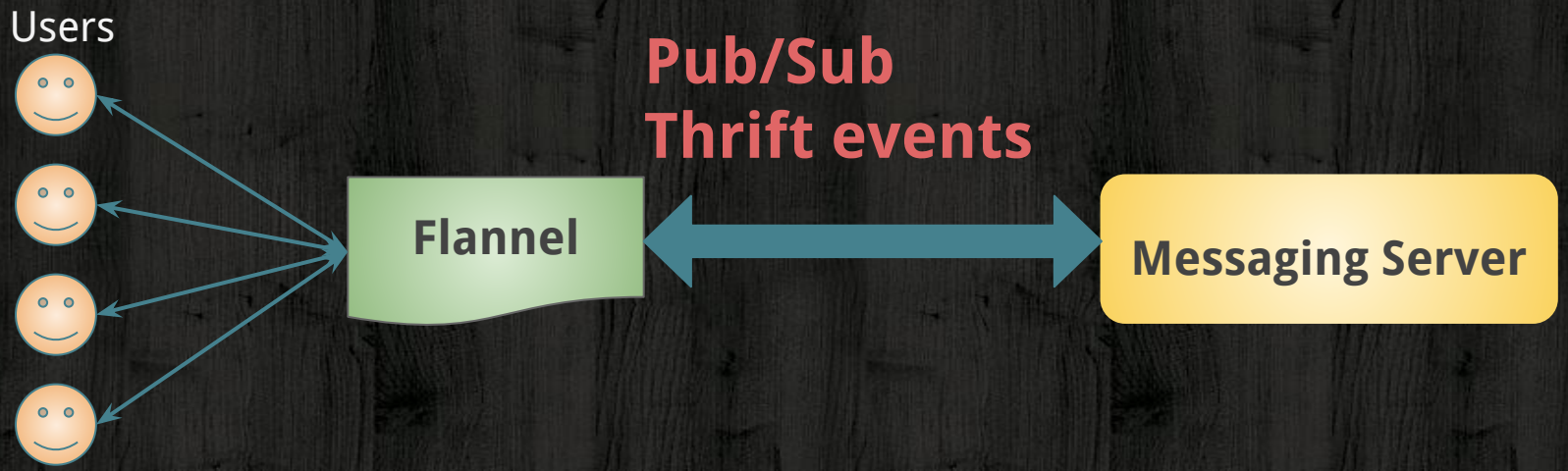
A Closer Look

Why does Flannel sit on WebSocket?

Old Way of Cache Updates



Publish/Subscribe (Pub/Sub) to Update Cache



Pub/Sub Benefits

Less Flannel CPU
Simpler Flannel code
Schema data
Flexibility for cache management

Flexibility for Cache Management

Previously

- Load when the first user connects
- Unload when the last user disconnects

Flexibility for Cache Management

With Pub/Sub

- Isolate received events from user connections

Another Closer Look

- With Pub/Sub, does Flannel need to be on WebSocket path?

Next Step

Move Flannel out of WebSocket path

Next Step

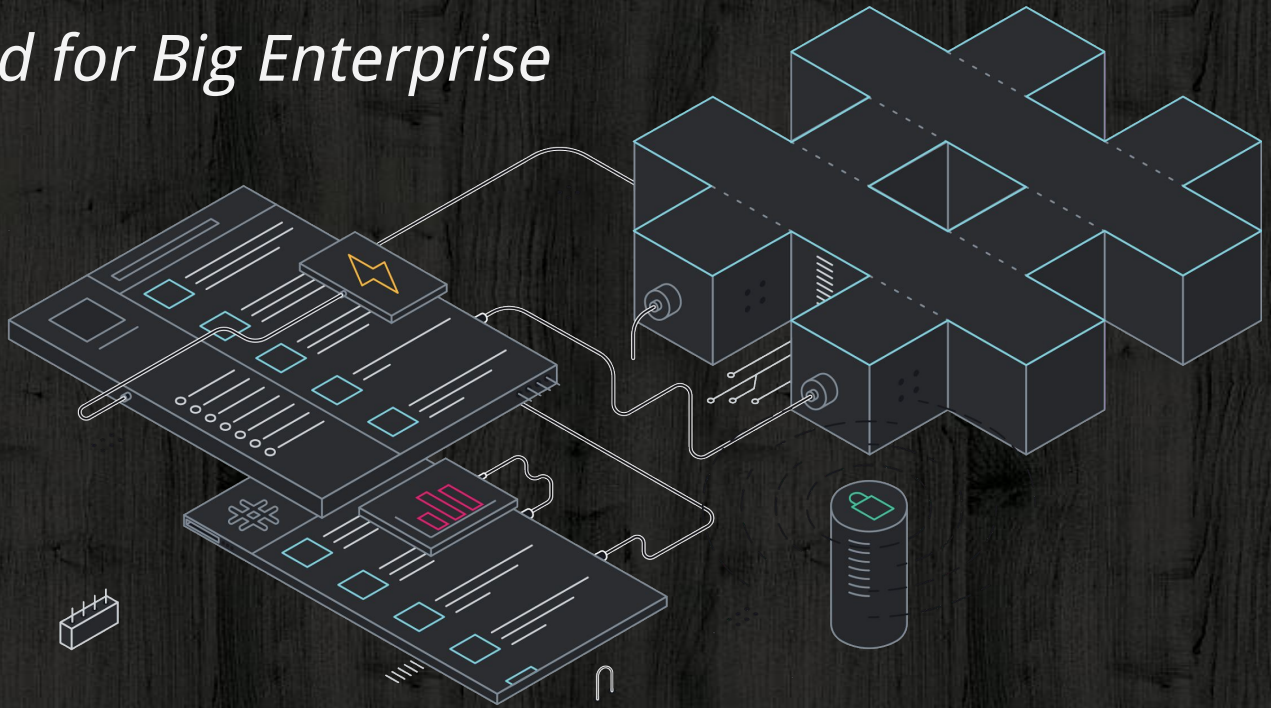
Move Flannel out of WebSocket path

Why?

Separation & Flexibility

Evolution with Product Requirements

Grid for Big Enterprise



Before Grid

Team Affinity for Cache Efficiency

Now

Team Affinity

Grid Aware

Grid Awareness Improvements

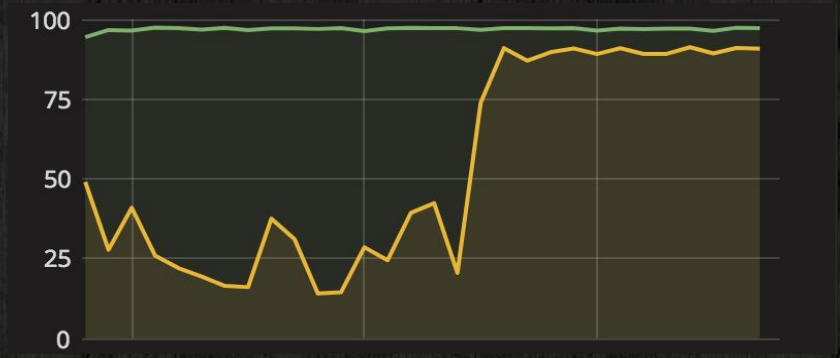
Flannel Memory

Saves 22G of per host, 1.1TB total

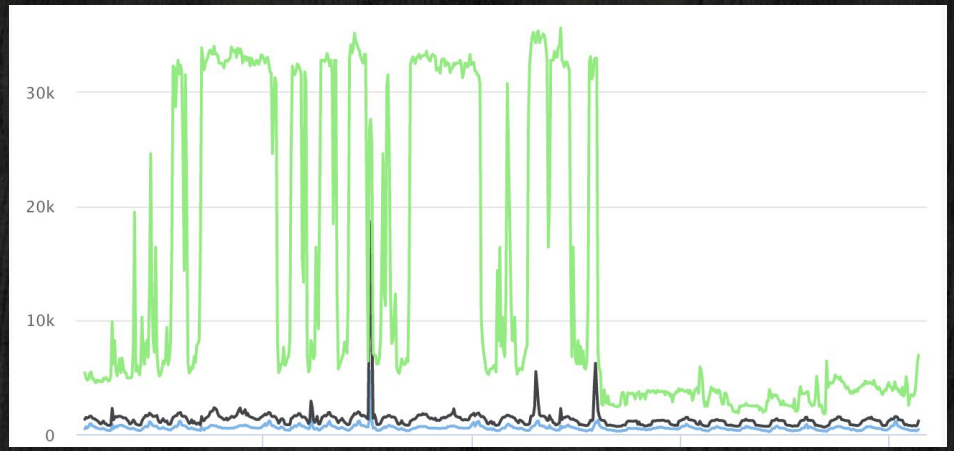
Grid Awareness Improvements

For our biggest customer

DB Shard CPU Idle 25% -> 90%



P99 User Connect Latency 40s -> 4s



Future

Team Affinity

Grid Aware

Scatter & Gather

Outline

- The slowness problem
- Incremental Improvements
- Architecture changes
  - Flannel
  - Client Pub/Sub
  - Messaging Server

Expand Pub/Sub to Client Side

Client Side Pub/Sub
reduces events Clients have to handle

Presence Events

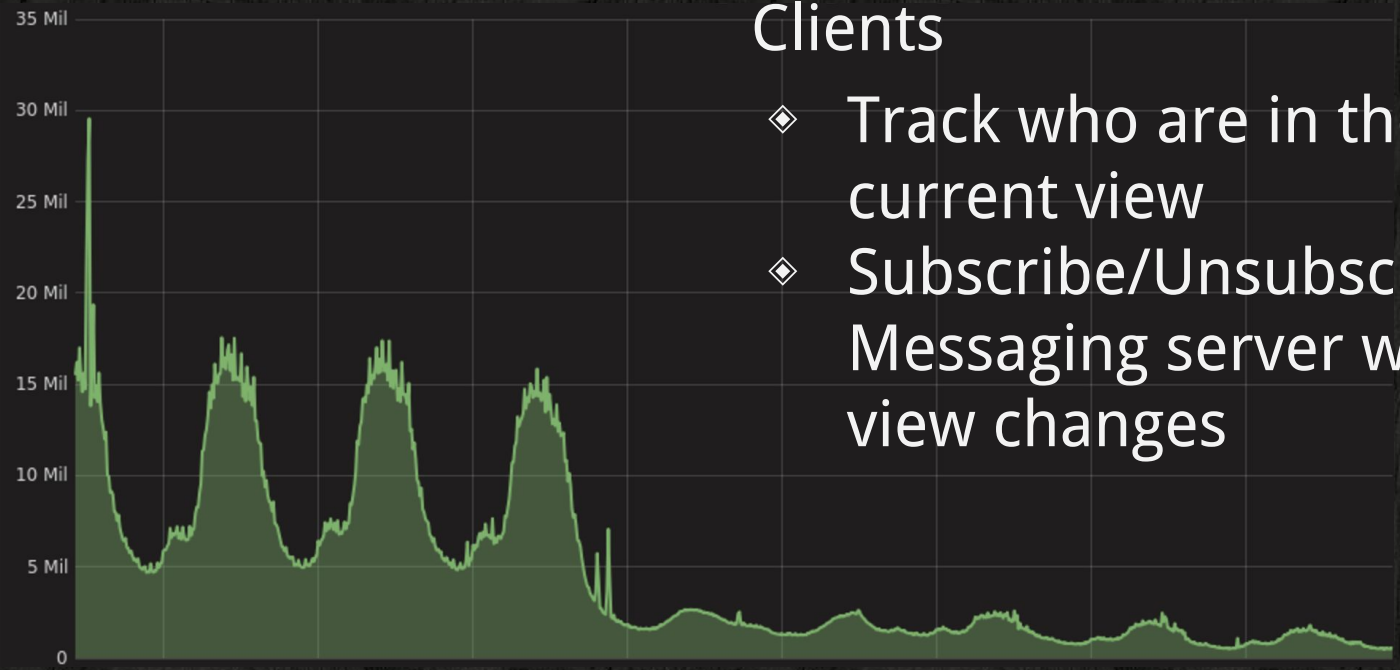
60% of all events

$O(N^2)$

1000 user team ⇒

1000 * 1000 = 1M events

Presence Pub/Sub



Clients

- Track who are in the current view
- Subscribe/Unsubscribe to Messaging server when view changes

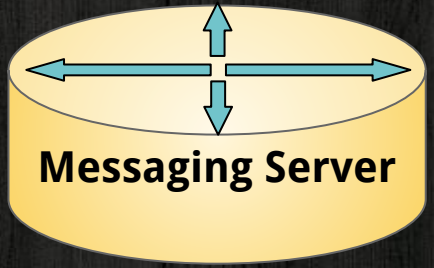
Outline

- The slowness problem
- Incremental Improvements
- Architecture changes
  - Flannel
  - Client Pub/Sub
  - Messaging Server

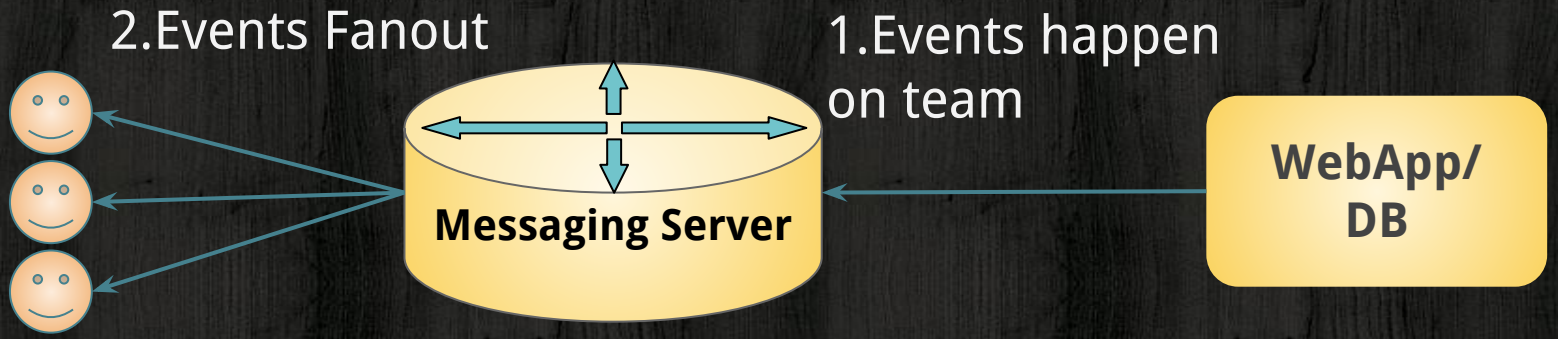
What is Messaging Server

Messaging Server

A Message Router



Events Routing and Fanout



Limitations

- Sharded by Team

Single point of failure

Limitations

- Sharded by Team

Single point of failure

- Shared Channels

Shared states among teams

PLAN A B

Topic Sharding

Everything is a Topic

public/private channel, DM, group DM, user, team, grid

Topic Sharding

Natural fit for shared channels

Topic Sharding

👍 Natural fit for shared channels

✅ Reduce user perceived failures

Other Improvements

Auto failure recovery

Other Improvements

Auto failure recovery
Publish/Subscribe

Other Improvements

Auto failure recovery
Publish/Subscribe
Fanout at the edge

Our Journey

Problem
Incremental Change
Architectural Change
Ongoing Evolution

Journey Ahead

Get in touch:
https://slack.com/jobs

Thanks!

Any questions?

@bingw11