

Google™



"Run corporate applications on Google App Engine? Yes we do."

Ben Fried, Google CIO

Justin McWilliams

Matt Simmons

Irwin Boutboul

2010-05-19

**View live notes and ask questions about
this session on Google Wave**

<http://bit.ly/appengine2>

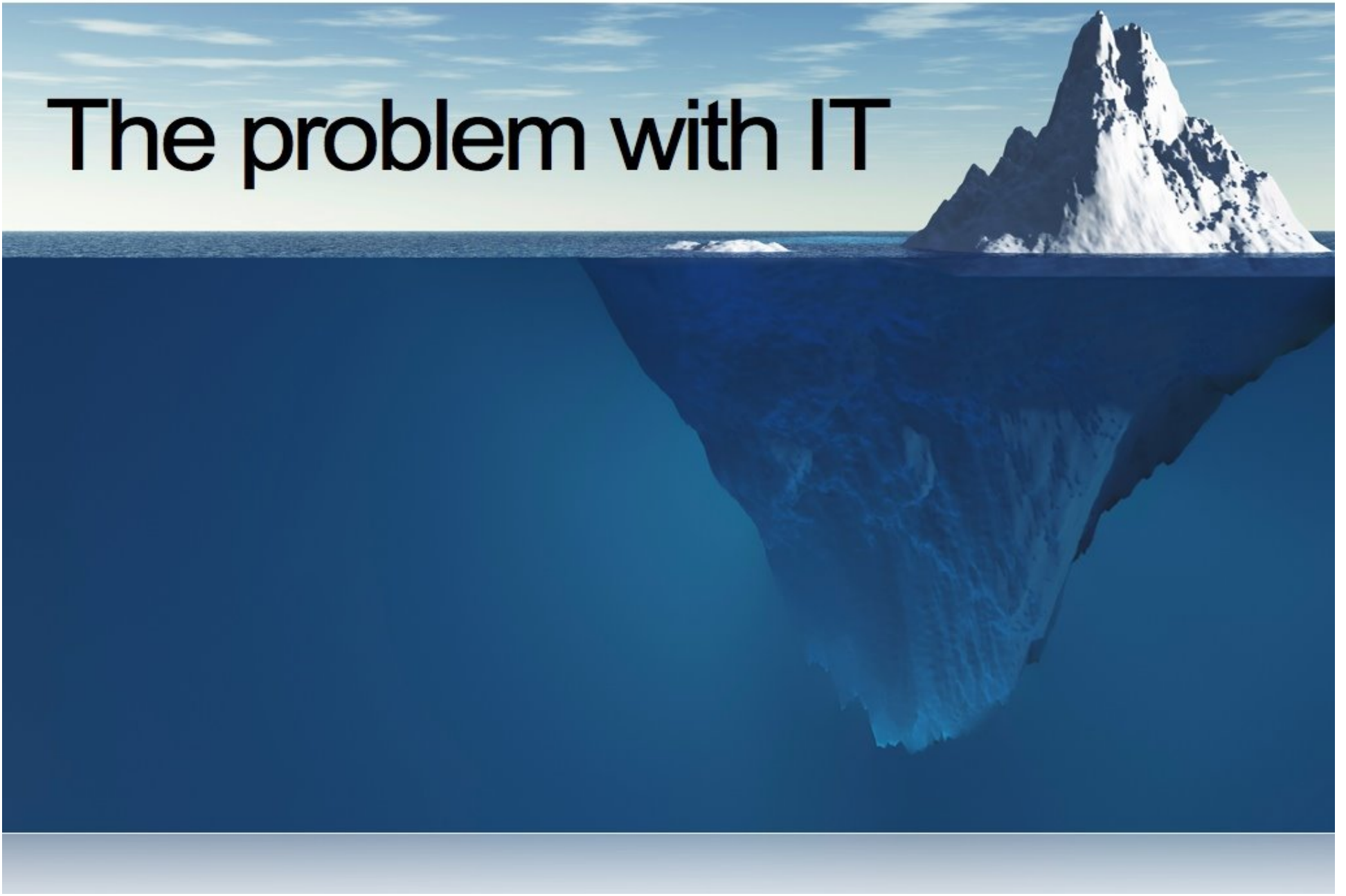
Introductions



What?

Why?

The problem with IT





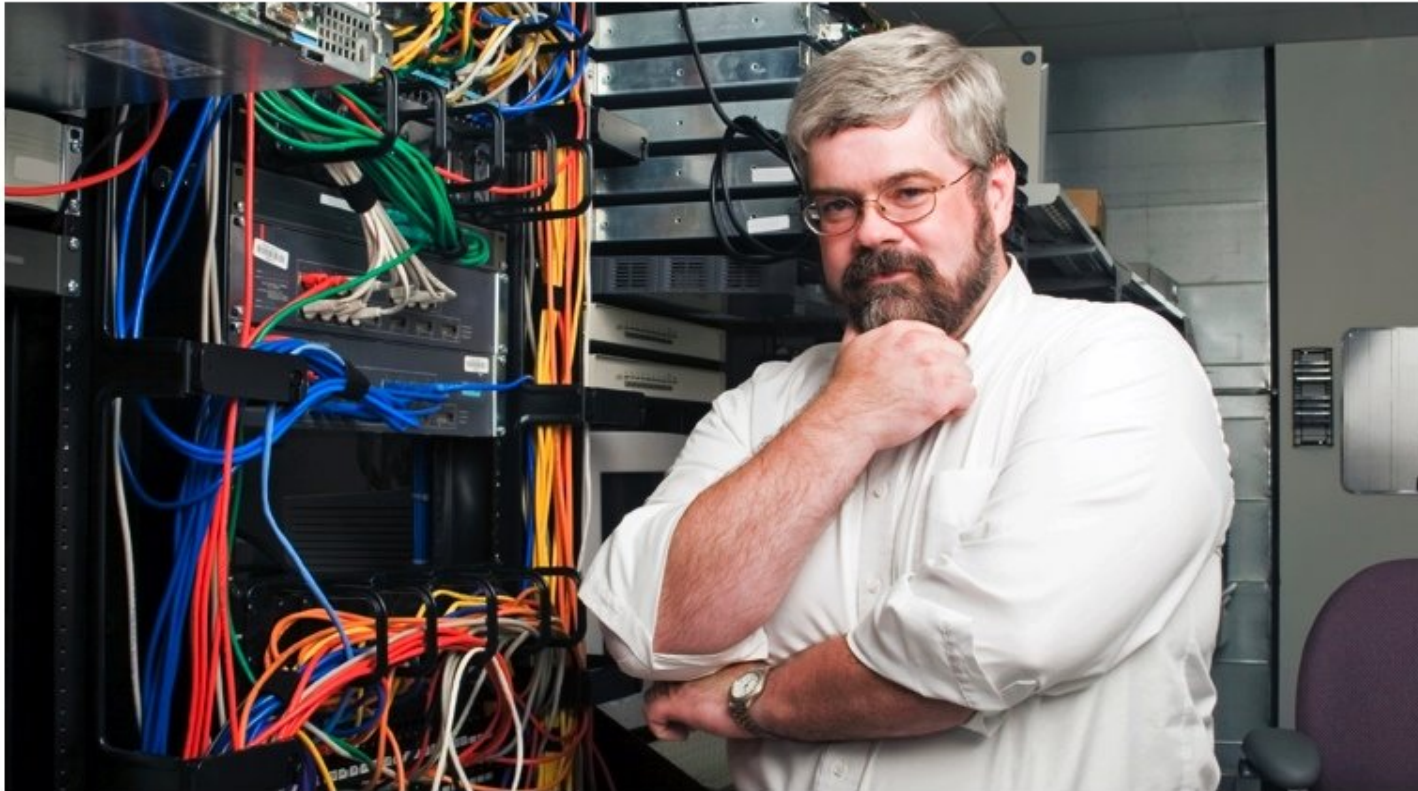




Just Say "No"

There's a better way

Have to do things differently



It's not easy

We've
figured it
out





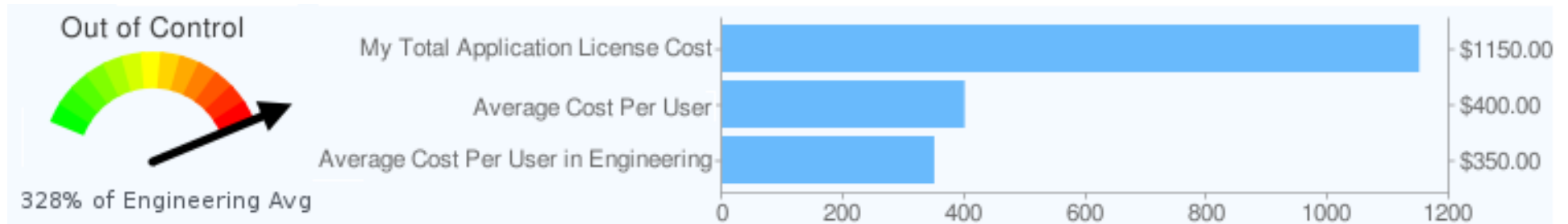
We'll show you how

Introducing AppReduce, Tech Stop Online, and CloudCourse

AppReduce

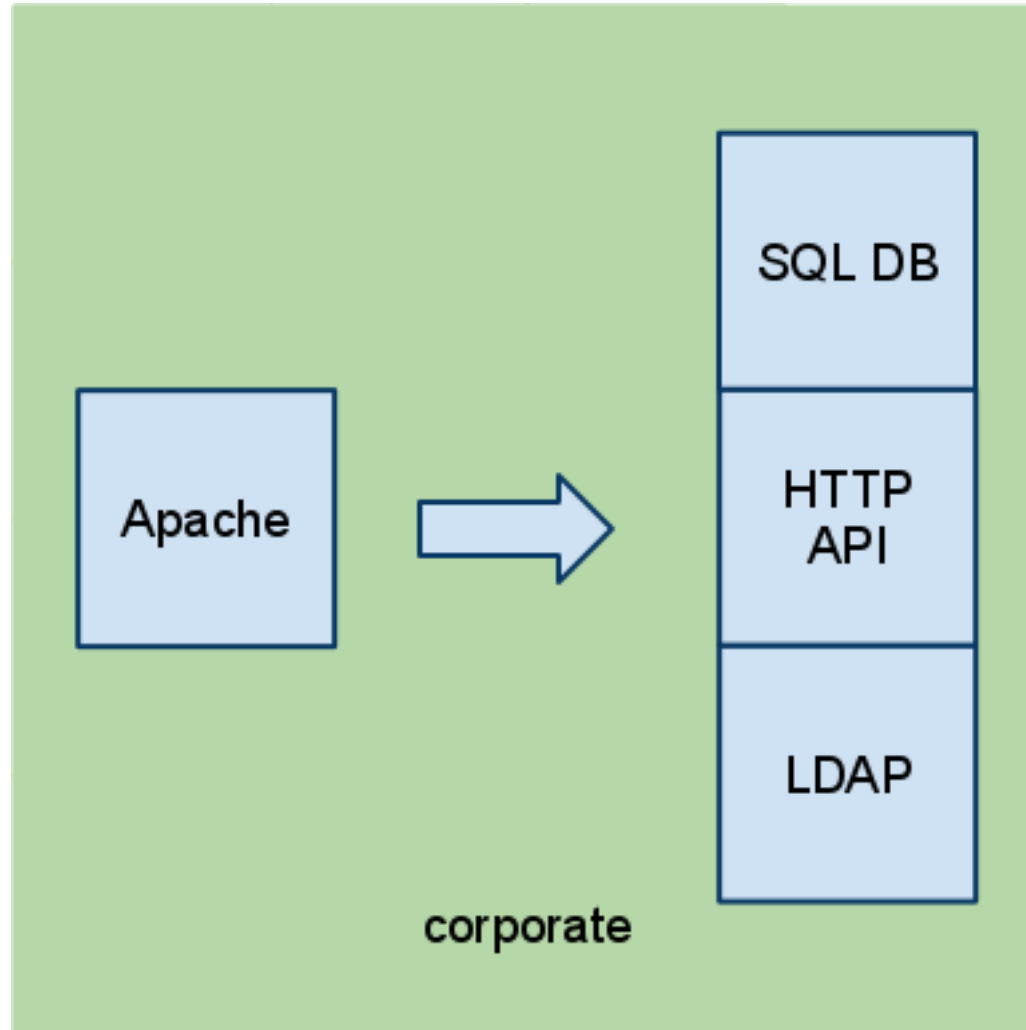
Justin McWilliams, CorpEng Developer

Reduce License Spend



Uninstall	Application	Cost	Alternative(s)
Uninstall	Foo Product 1	\$200.00	View
Uninstall	Foo Product 2	\$350.00	View
Total Application Cost		\$550.00	

Design Iterations - LAMP



Framework Familiarities

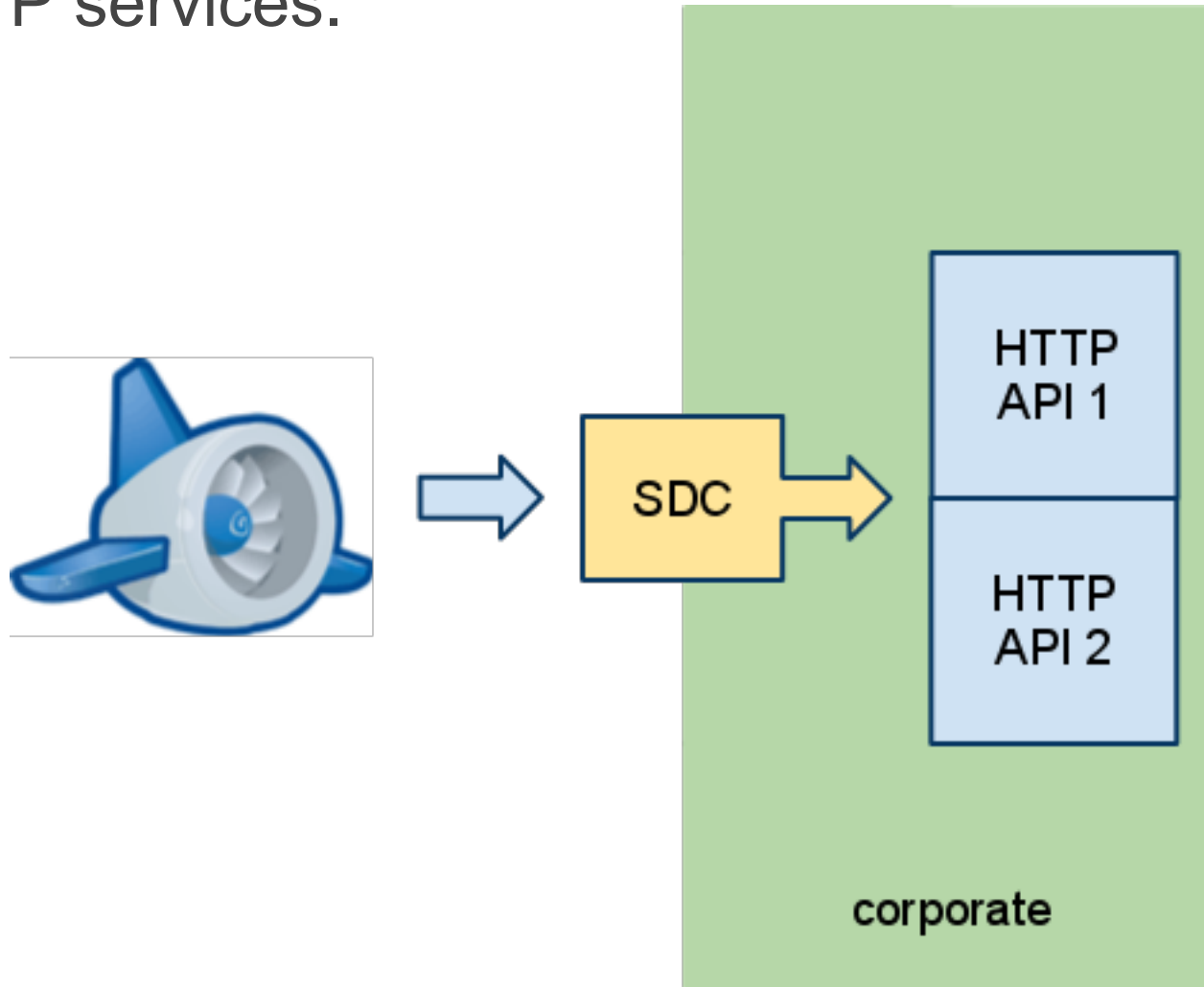
- Python
- Django
- Google Closure JavaScript Library
- Google Chart API
- Google Analytics

Challenges

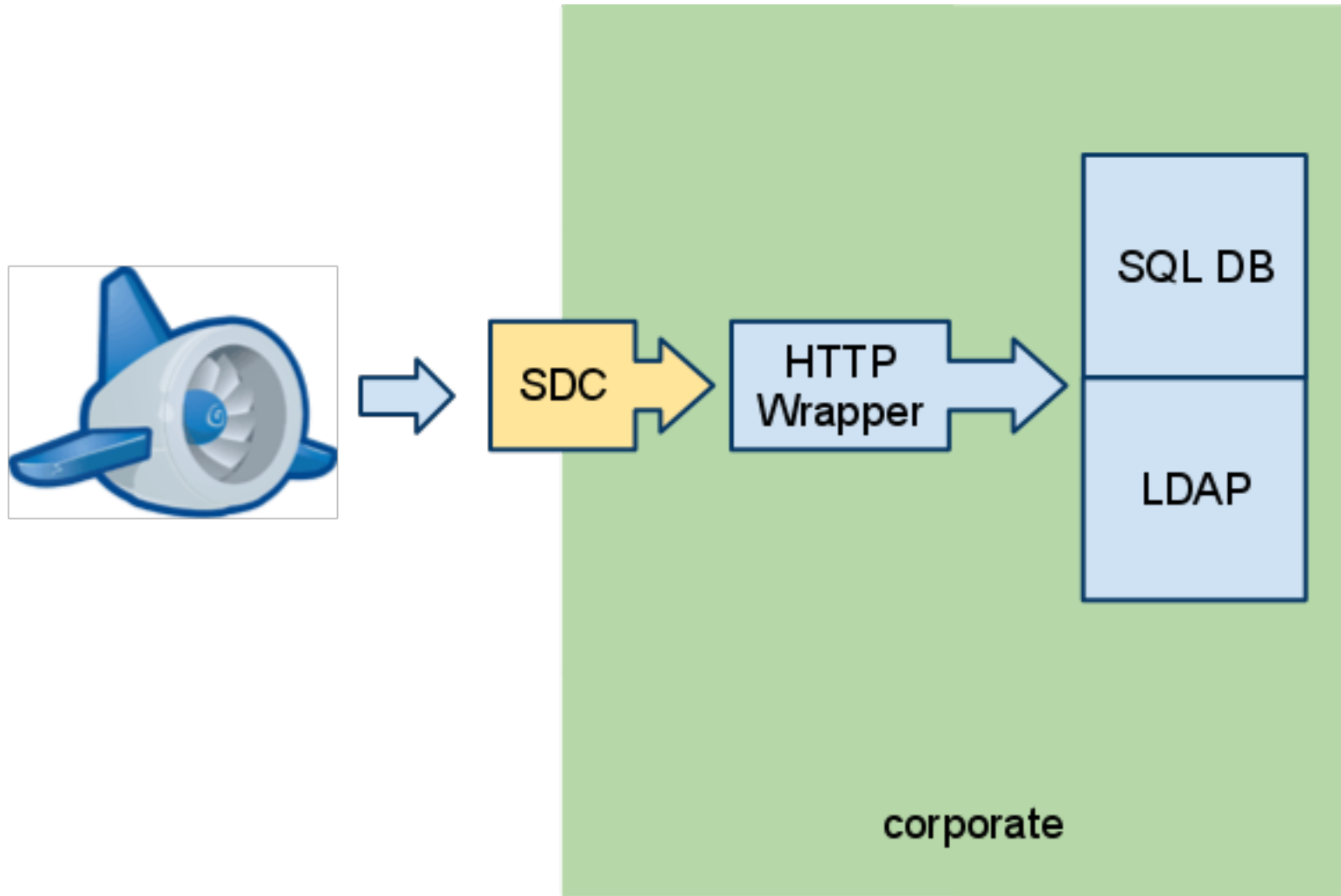
- **Accessing corp APIs and services**

SDC Background

Secure Data Connector (aka SDC) allows for a secure tunnel to Corp HTTP services.



SDC: Wrap non-HTTP services

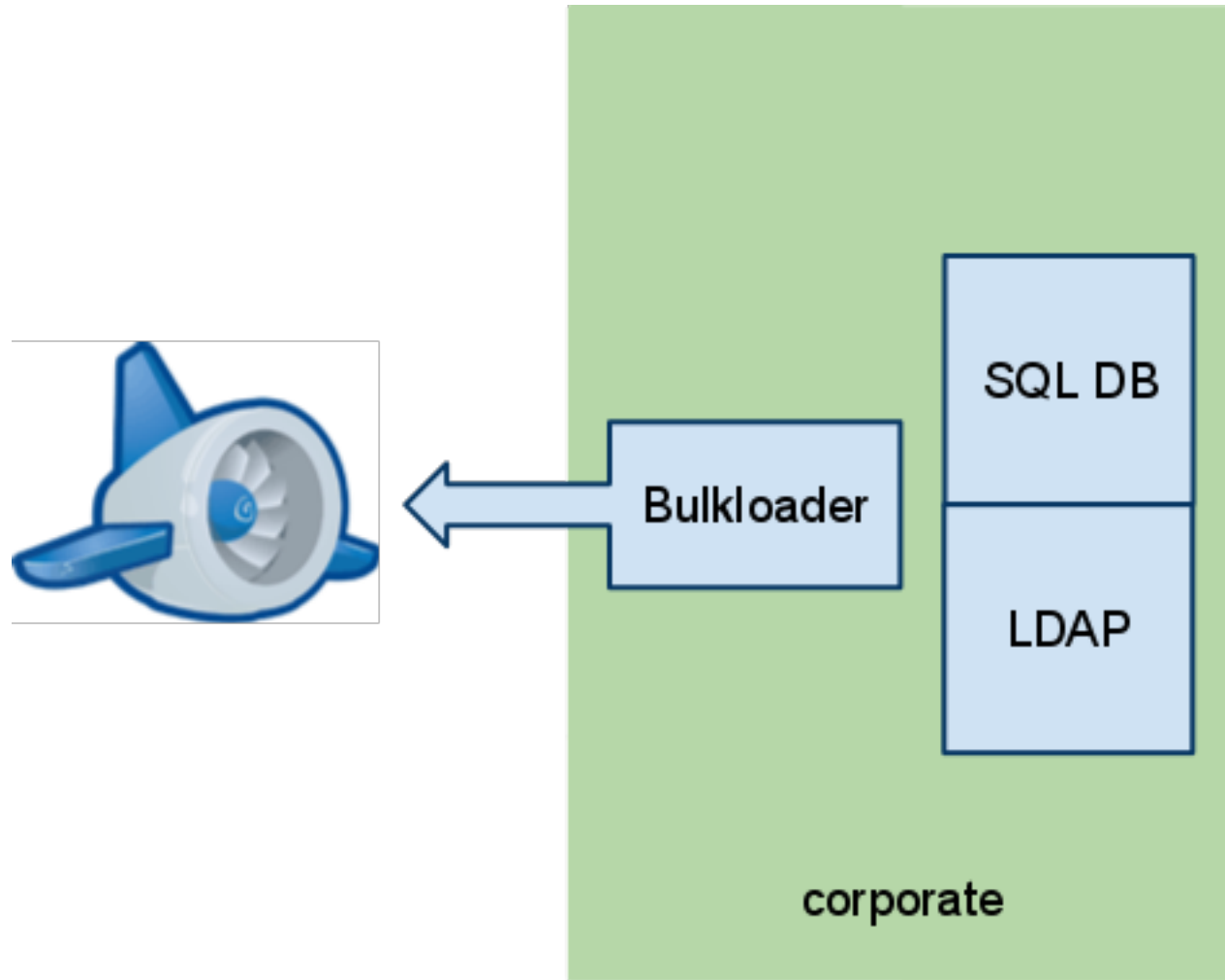


Challenges

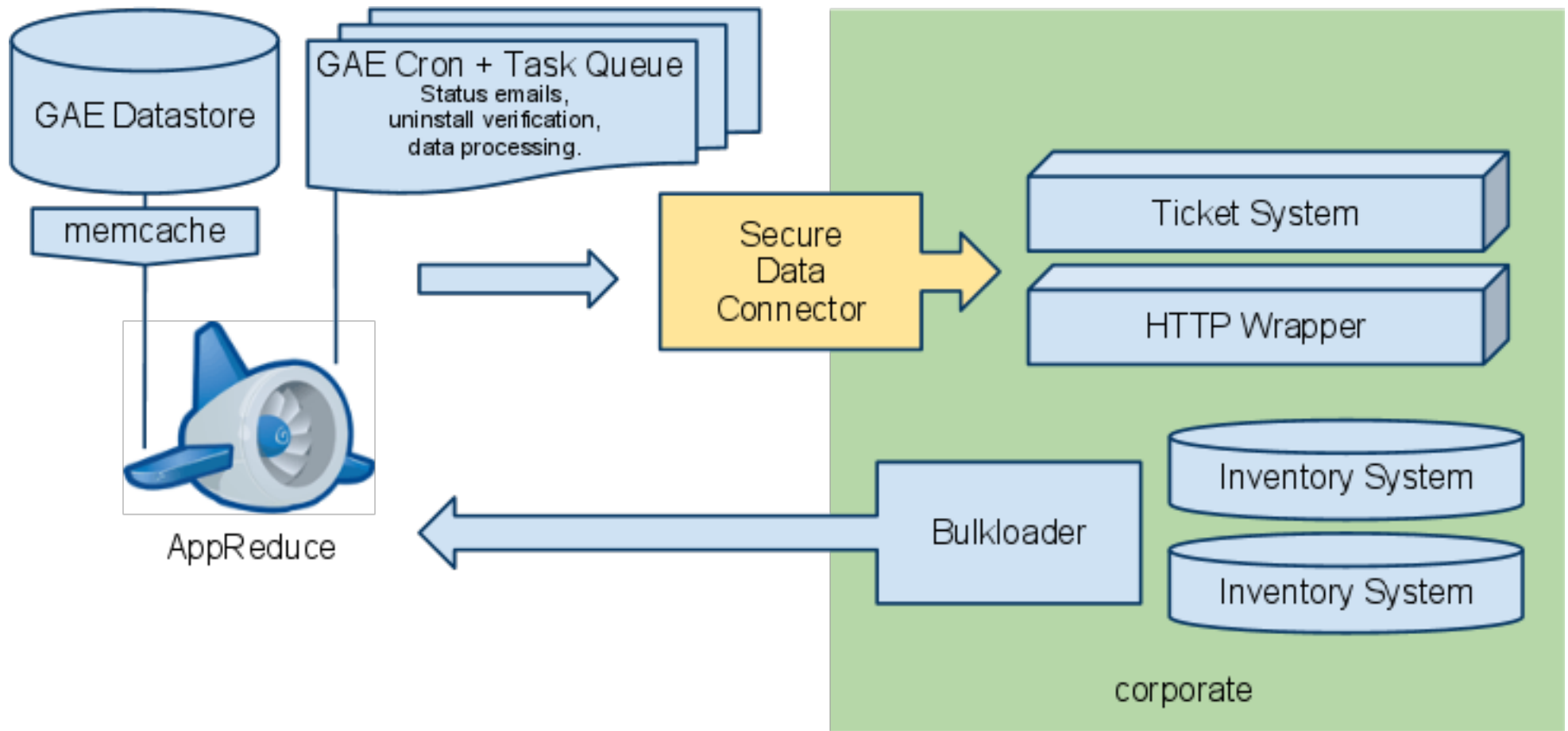
- Accessing corp APIs and services
- **Realtime user performance**

Cache corp data with Bulkloader

Bulkloader makes it easy to import large amounts of data from anywhere into AppEngine.



Design Iterations - GAE



Results and Success

- Putting information into employees' hands can be powerful
- \$450K+ worth of returns to date
- \$150K+ in first 24 hours after launch
 - 50 QPS spike; 20 QPS sustained
- Built, released and supported by only 2 engineers

TechStop Online

Matthew Simmons, CorpEng Developer

Tech Stop Online: What's in a name?

- Home of internal Google IT support
- “Self-powered”



Tech Stop Online

- Documentation repository
- Managed LAMP -> App Engine
- Django + App Engine + Closure
- SDC, Task Queues, Bulk Loader

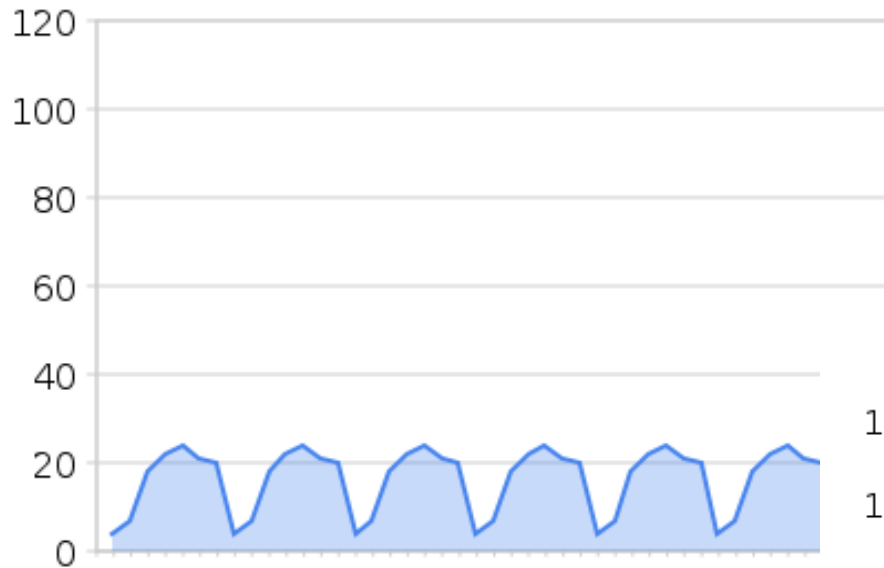
Motivations and Challenges

- Scalability
- Search
- Long running tasks



What is "Bursty" demand?

Normal Demand



"Bursty" Demand



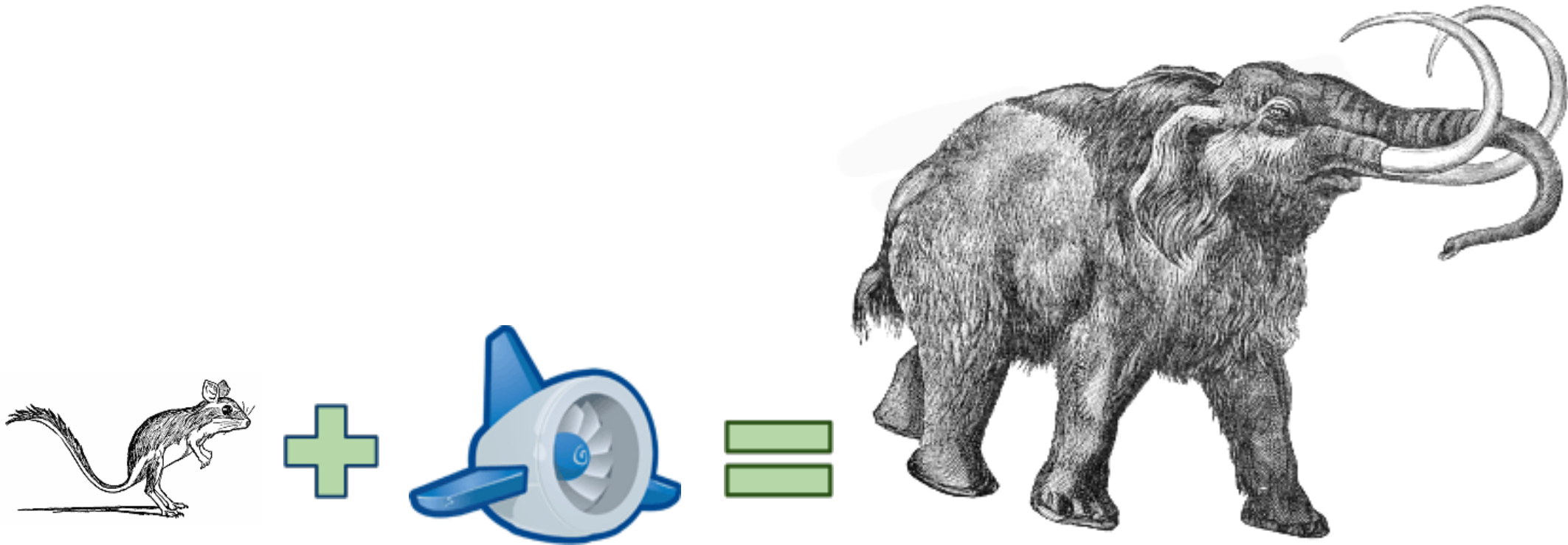
Scalability

- Old platform:
Scalability == Expensive
- Cycles from application development "stolen" by platform concerns.



AppEngine: Scalable by default

- It just works
- What did we do with our extra time?



Tech Stop Online: Navigation is king.

The Google logo is displayed in its characteristic multi-colored font (blue, red, yellow, blue, green, red).

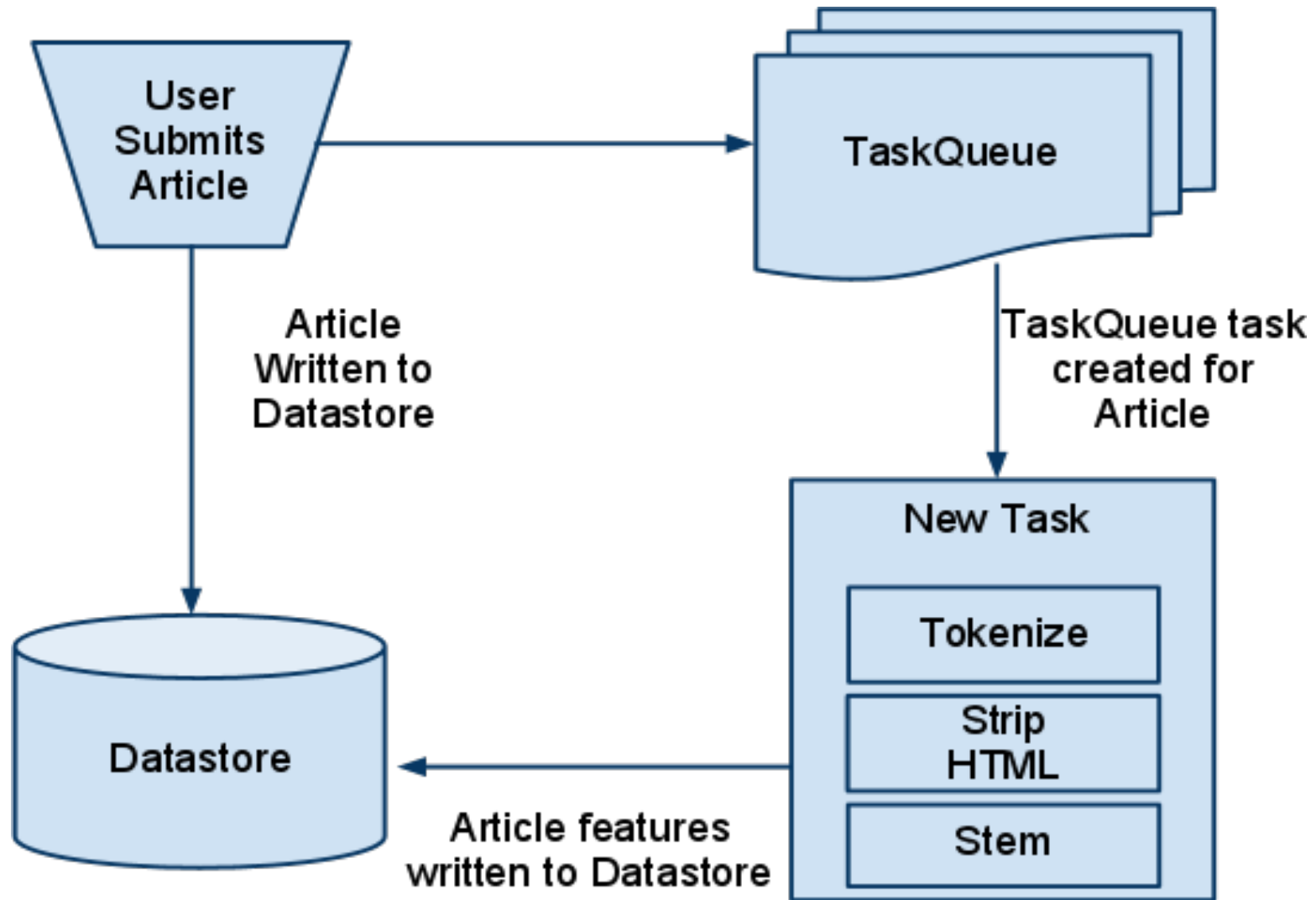
My computer is broken :(

Google Search

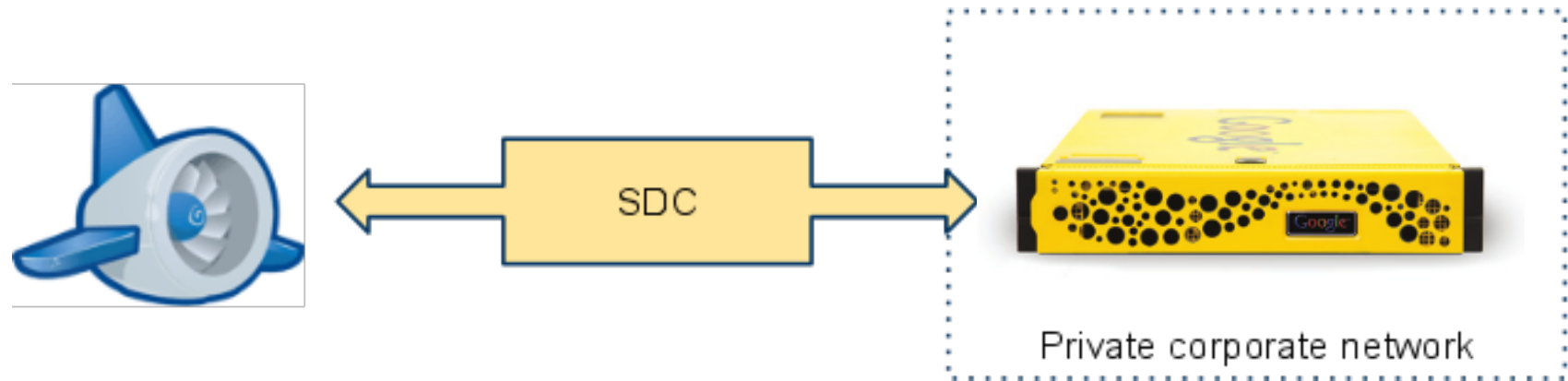
I'm Feeling Lucky

- Problem: No native TextProperty indexing
- Solutions
 - DIY indexing
 - Offload to Google Search Appliance

Search - Solution 1: DIY Indexing



Search - Solution 2: GSA Indexing



- Indexing

- Problem: content behind authentication
- Push feed to GSA index over SDC

- Searching

- Pull from GSA API over SDC
- Display results inline

Using SDC

```
from google.appengine.api import urlfetch

result = urlfetch.fetch(
    url='http://corp.example.com/people.csv',
    headers={'use_intranet': 'yes'})

if result.status_code == 200:
    parseCSV(result.content)
```

[Secure Data Connector - http://code.google.com/securedataconnector](http://code.google.com/securedataconnector)

Computing Similarity

- Goal: Automatically generate “related articles”
- Idea: Use document x document similarity
- Feasible, but...
 - New documents need to be added
 - Updates too much for a single request

Task Queues

- Defer expensive jobs to Task Queue
- Write a handler like usual, but run it asynchronously
- Gets the technology out of the way of the task.

Flatten Your Data

- Kiss BCNF goodbye!
- Datastore denormalization
- Use Reference properties where necessary
- Watch out for reference properties in loops
- It's OK to create "cached" copies of data on multiple models

Testing Instances: The Easy Way

```
application: techstop-online  
version: 1  
runtime: python  
api_version: 1
```

```
handlers:  
- url: /  
script: main.py
```

```
application: techstop-online-ga  
version: 1  
runtime: python  
api_version: 1
```

```
handlers:  
- url: /  
script: main.py
```

Testing Instances: Another Easy Way

```
application: techstop-online-qa  
version: matt  
runtime: python  
api_version: 1
```

```
handlers:  
- url: /  
script: main.py
```

```
application: techstop-online-qa  
version: justin  
runtime: python  
api_version: 1
```

```
handlers:  
- url: /  
script: main.py
```

Summary

- Scalable by default!
- Fetch corporate data with SDC
- Defer expensive jobs to Task Queue
- Datastore denormalization
- Set up a -qa instance and iterate away

CloudCourse

Irwin Boutboul, CorpEng Developer

1 Create an Activity

Create new activity

Title

Activity type In-person Video Virtual classroom/webinar Assesment or test

Description [Link](#) **B** *I* Normal / serif Large [Edit HTML](#)

An introduction to programming in Python

Difficulty optional

2 Schedule a Session



Python 101 > Schedule a session

General information

Maximum registration optional

Registration deadline optional

Reserve rooms

Adding timeslots

When to [America/Los_Angeles](#)

Primary location MTV-1055-1-Beka (5) VC

Other locations optional CHI-KIN-7-Metro Chicago (12) VC

VC bridge optional

3 Find a session

Activity

Python 101

☆☆☆☆☆

0 Ratings

☆☆☆☆☆

No rating

Description

Introduction to programming in Python.

Activity type: In-person

Level: Introductory

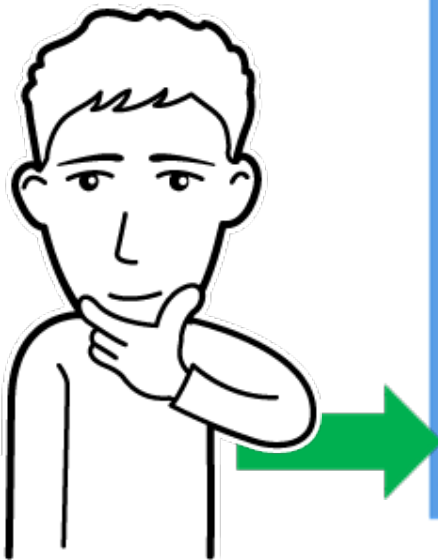
Contact: [python-class](#)

Offered by: EngEDU

Restricted to: Employee, Intern

Sessions

Day	Date	Time	Location	Instructor	Status	
Sun	Apr 25, 2010	2:00am - 8:00am PDT	London	sharadha	FULL (waitlisted: 2)	<input type="button" value="Register"/> link user edit trash
Tue	Apr 27, 2010	2:53pm - 3:53pm PDT	Hyderabad	srid	OPEN (remaining: 5)	<input type="button" value="Register"/> link user edit trash
Mon	Jun 7, 2010	2:00pm - 4:00pm PDT	Mountain View 1 more »	srid	OPEN (remaining: 1)	<input type="button" value="Register"/> link user edit trash



Technology

- App Engine / Python
- Django (w/ patch)
- Closure
- GData

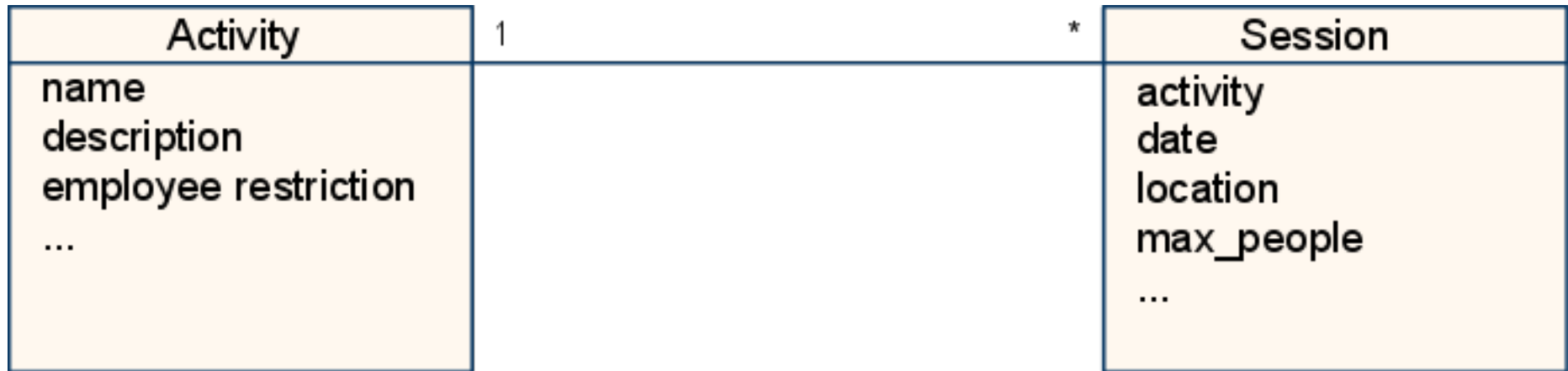
Challenges

- No joins
- Long running computation
- High performance transactional system

Challenges

- **No joins**
- Long running computation
- High performance transactional system

Schema - Relational



Python 101

June 5th - MTV
June 6th - SFO
June 6th - Webcast

Schema - non relational

Activity
name
[session1 session2... sessionN]
[date1:location1 ... dateN:locationN]
...

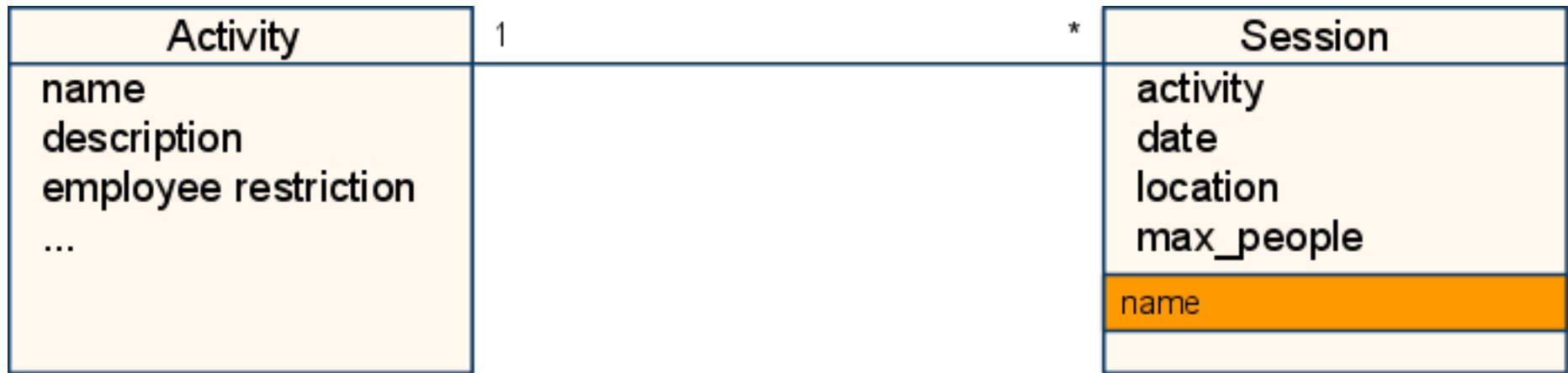
Python 101

[session1 sessionN]

[June 5th:MTV, ..., June 6th:VC]

- exploding indices
- 1 MB limit
- performance / quota

Schema - non relational



- exploding indexes
- 1 MB limit
- performance / quota
- trade-offs

Challenges

- No joins
- **Long running computation**
- High performance transactional system

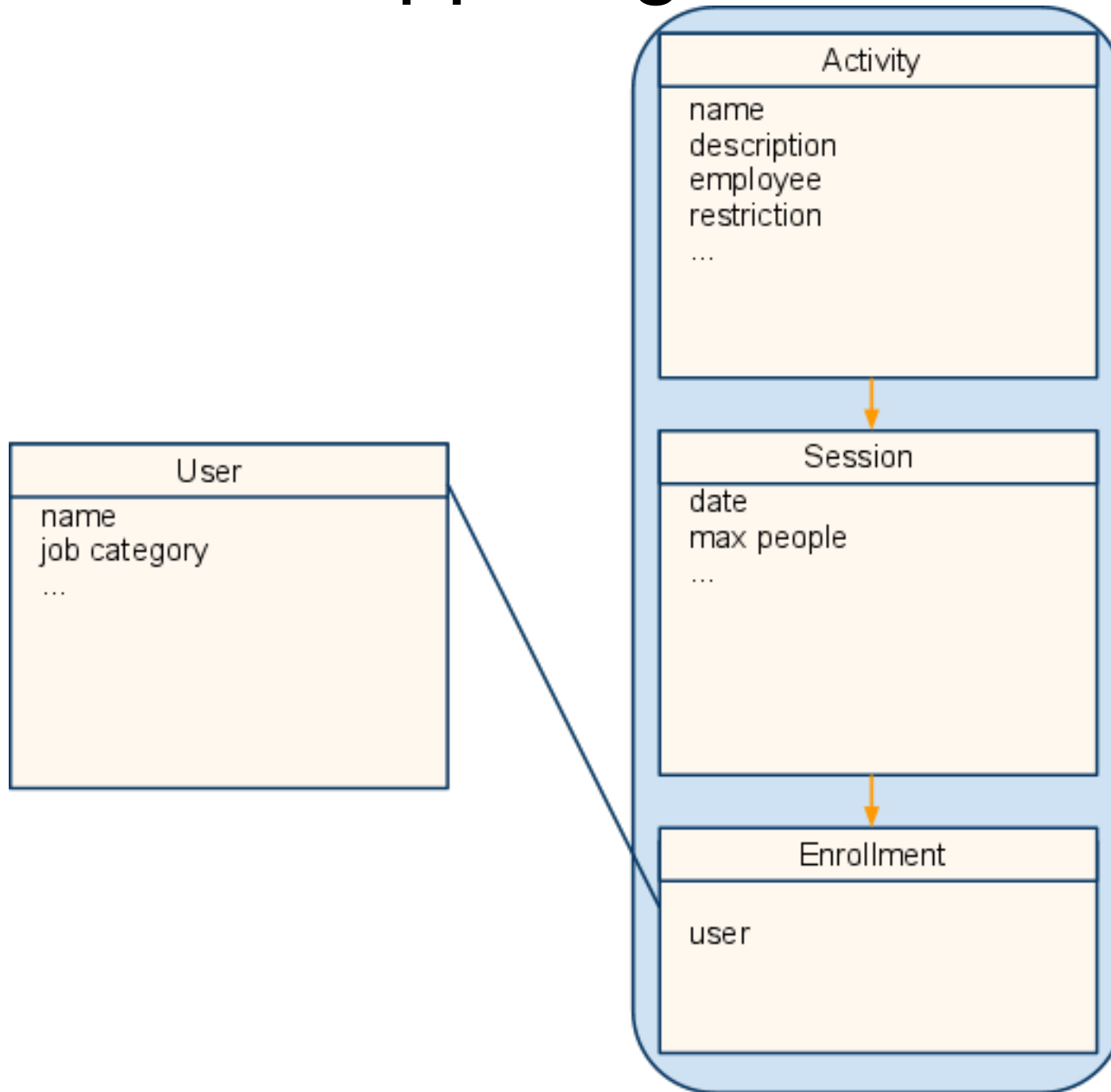
Long running computation

- Push burden of work to backend processes
- Defer tasks / transactional
- Tasks chaining / cron jobs

Challenges

- No joins
- Long running computation
- **High performance transactional system**

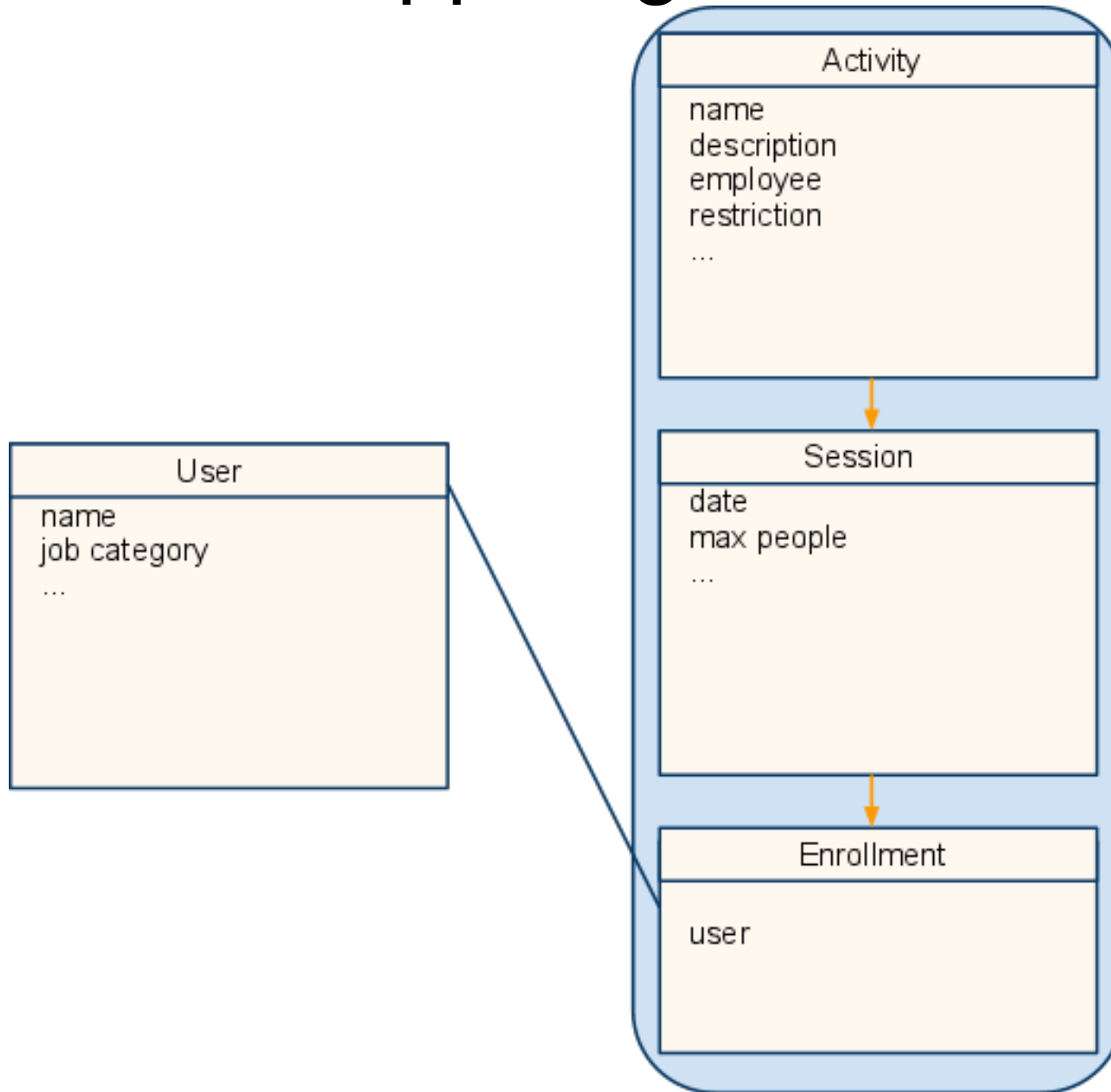
Schema - App Engine



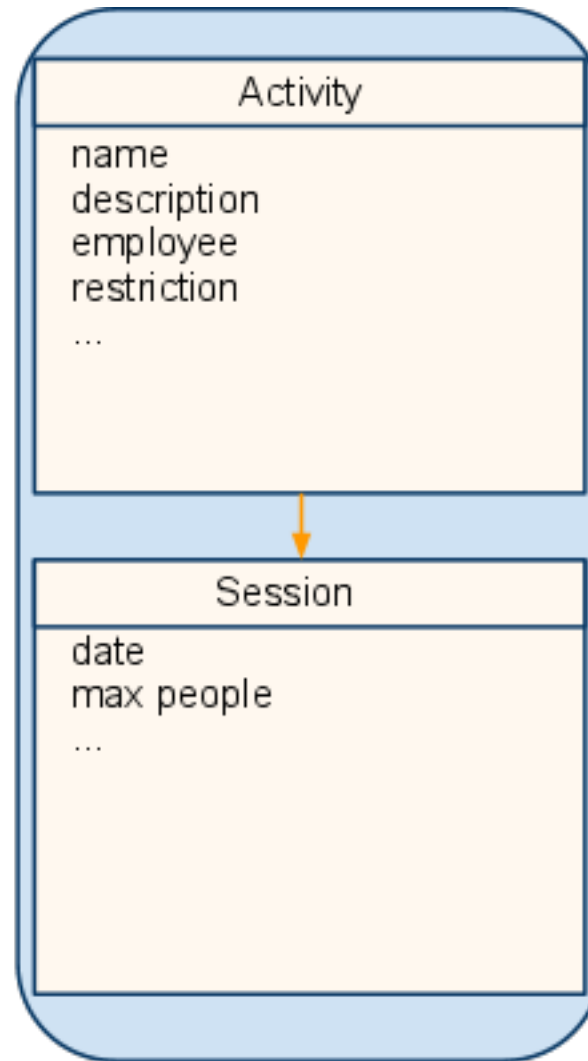
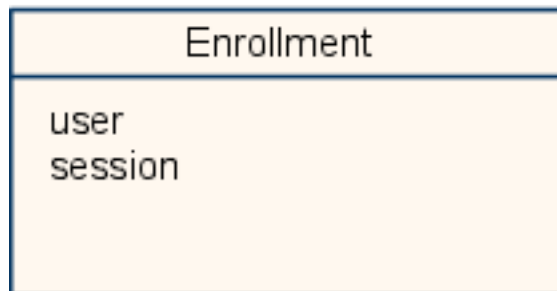
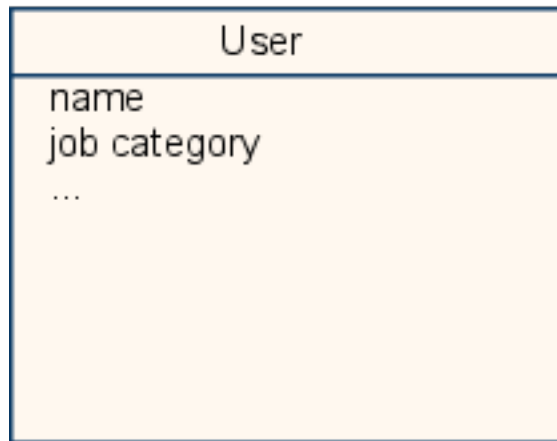
Enterprise constraints

- Each signup is a transaction on an activity
- 1-10 writes/sec per entity group
- Too slow !

Schema - App Engine



Schema - App Engine

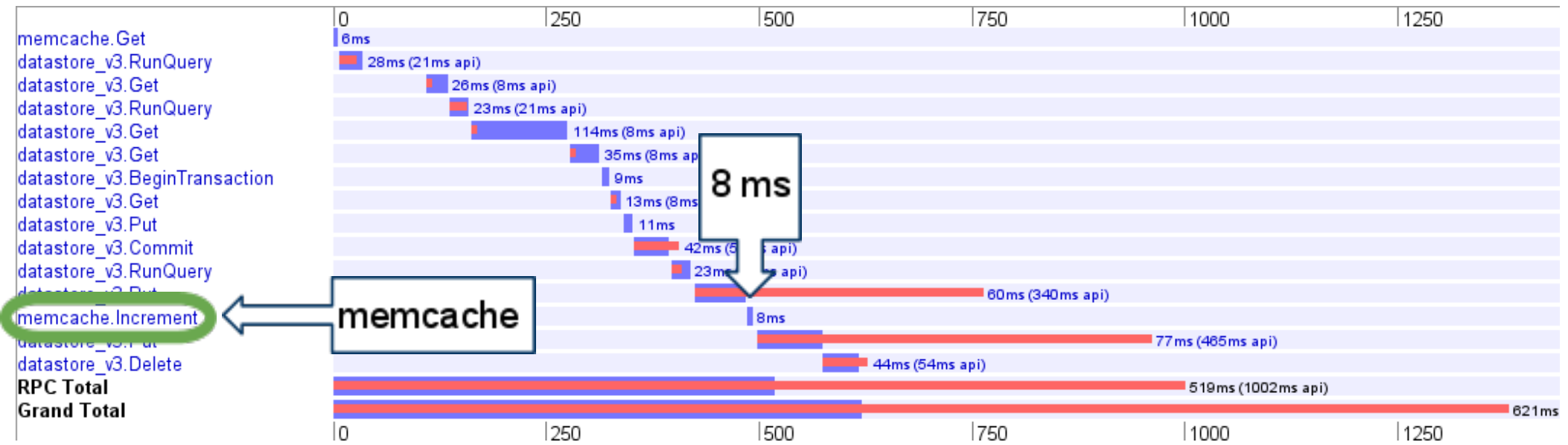


Solution

- `memcache.incr(max_people)`
- Conservative response
- High level locks offline -> correct response

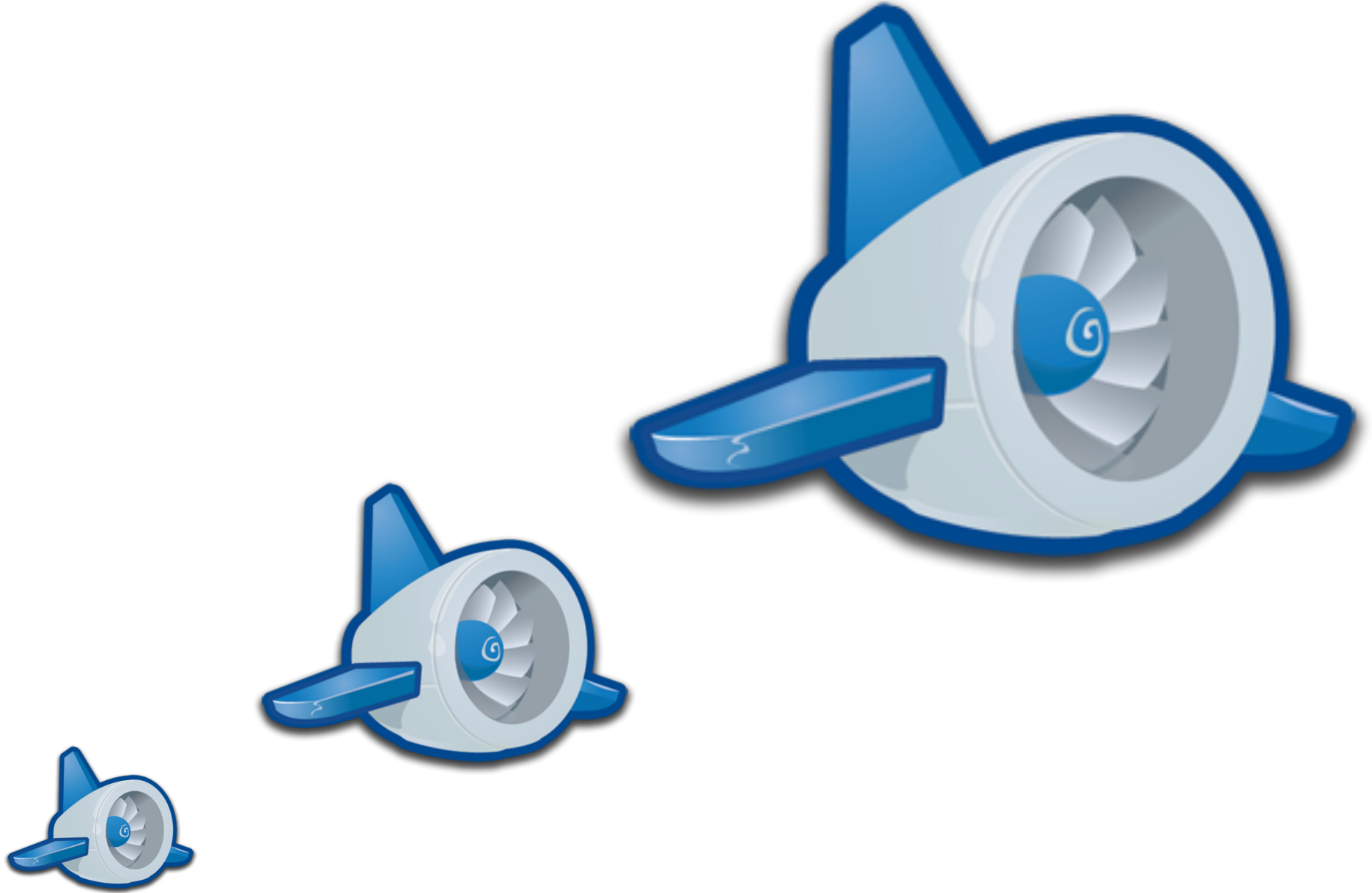
Scalability

Online ~ 600 ms per enrollment



100+ writes per second for an activity

Scalability



Summary

- De-normalization trade-offs
- Asynchronous processing
- Leverage memcache atomic ops

Releases and Sandbox Demos

CloudCourse

- <http://code.google.com/p/cloudcourse>

AppReduce

- <http://code.google.com/p/appreduce>
- <http://code.google.com/p/gord>

Techstop Online: coming soon

Visit our sandbox booth for demos and questions!

View live notes and ask questions about this session on Google Wave

<http://bit.ly/appengine2>

Google™

