Avoiding Social Lock-In with Crosscloud

Sandro Hawke, W3C / MIT, sandro@hawke.org
W3C Tech Talk
5 Jun 2014
http://www.w3.org/2014/Talks/0605-sandro

Overview

- The Problem
  - Social Lock-In
- The Solution
  - Crosscloud = Blogs + RDF + Auth + ...
- The Plan
  - Research
  - Community
  - Standards
  - Business

Product Lock-In

Once you start using a product, there can be a high cost to switch. This is old news.

Most common: Data Format Lock-In

- .doc files (“proprietary formats”)
- ever switch accounting systems?

Mitigated by standards, non-binary formats

Social (Multi-User) Software

You're interacting with people via computer

- chat
- auction
- meeting scheduling
- dating
- video sharing
- … many more

The value is largely in **other people** using it.

Social Lock-In

If you switch, you find yourself in a ghost town, unless you can get your primary contacts to move at the same time.

Re diaspora (facebook competitor): “It’s cool, but the sad fact is most people will still go to Facebook simply because that's where **all their friends** are." -- Newsweek

It’s not insurmountable (cf snapchat, g+) but it is very strong.

Examples

Imagine these without other users:

eBay, airbnb, yelp, foursquare, kickstarter, evite, meetup, doodle, eHarmony, okcupid, digg, reddit, facebook, linkedin, twitter, tumblr, instagram, pinterest, flickr, wikipedia, StackOverflow, Words with Friends, Github, ...

Very high barrier to entry for competition

Results?

Other developers (innovation) are locked out.

Little “serious” (enterprise) use.

Exclusion. Not everyone’s on every system.

No real options for privacy.

The systems change when *they* want.

So what is Crosscloud?

A system (under construction) that doesn’t have these problems.

You can switch apps, still working with the same people and data

Healthy competition:

- vi/emacs
- firefox/chrome
- elm/pine/thunderbird

Great for Free/Open Source Software

As a coder, even if github's software were 100% free, I couldn't effectively fork it (social lockin, all the other users)

With Crosscloud I could.

Don't like the latest change? Just keep running the older version of the code.

Share your own changes with the world

Of course it's Decentralized

We're building it, but once it's out there, no one will be able to control it.

- Like the Web
- Like email

People can join & participate without asking anyone's permission or paying anyone (beyond various service providers, which could be ad-subsidized).

- Good for weekend hackers *and* for enterprise

Problem Summary

Social Lock-in comes naturally with Social Software (where the value is the people)

It locks out innovation, business applications, government applications

It forces people to accept random site changes, privacy invasion, ads, questionable security

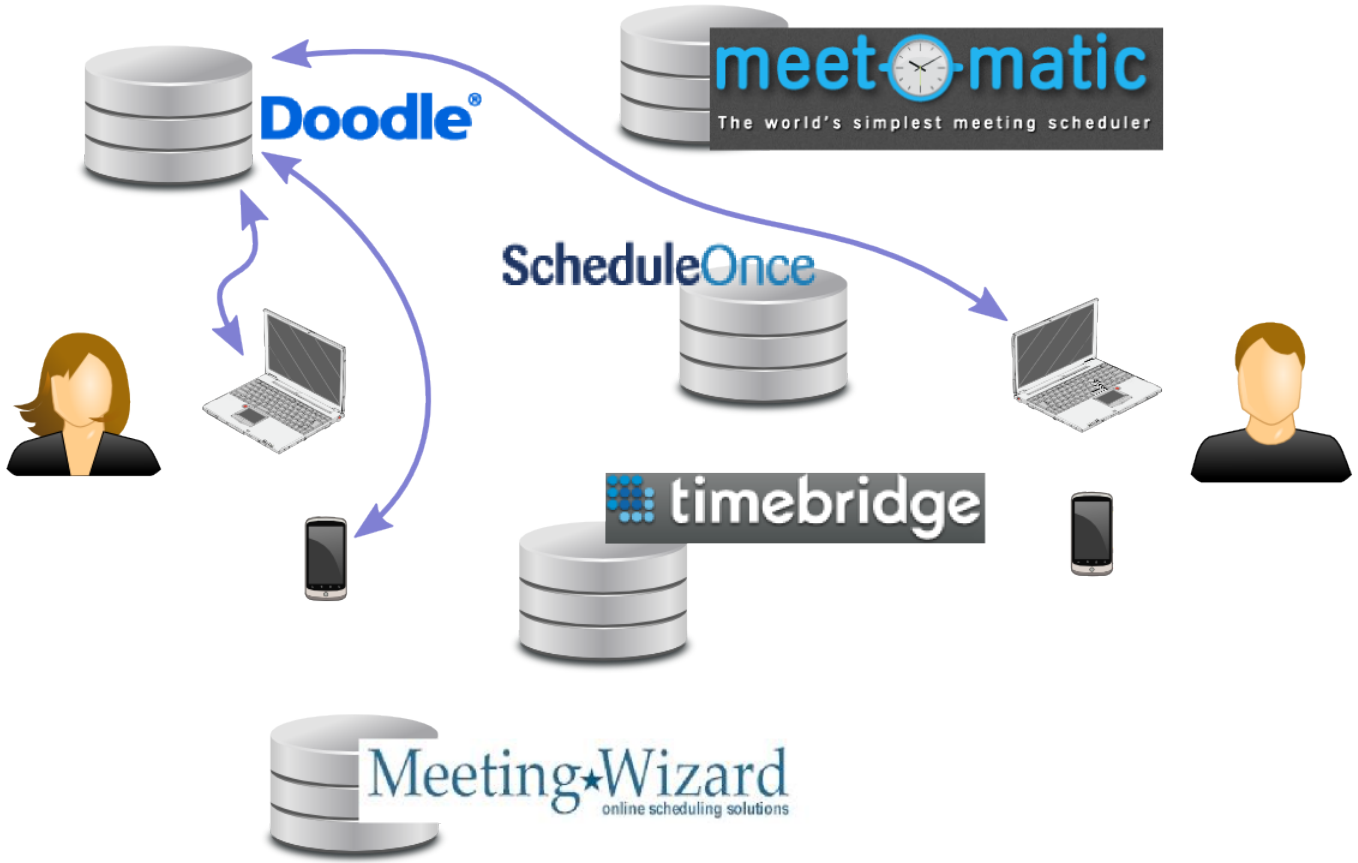
We imagine Crosscloud as **a platform without Lock-in**.

So how could this possibly work?

It’s about the data.

Meeting Scheduling Today

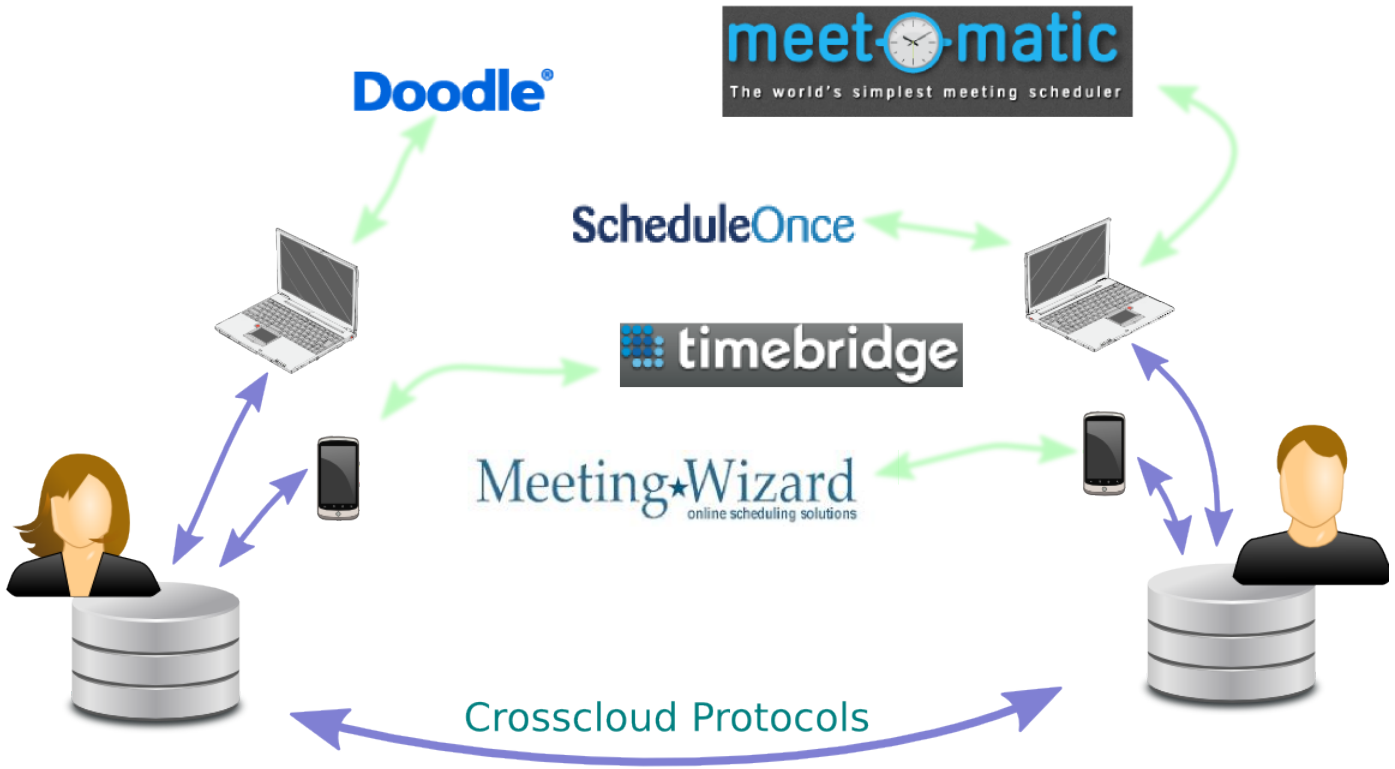
Scheduling a meeting today (without Crosscloud)



Everyone invited has to use the same vendor!

Meeting Scheduling with Crosscloud

Scheduling a meeting -- With Crosscloud



You remain in control of your data.
You are free to pick different software and service providers.

The Solution:

Blogs + RDF + Auth + ...

Blog =

- everyone has “their own” website
- link notification enables conversation

RDF =

- general data format
- anyone can extend to any problem domain

Auth =

- refer to people
- access control

Blog instead of Twitter

blogging vs microblogging? it's mostly UI.

tweeting = posting to your own microblog
following = your reader-software gathers it for you

Just post/store data, let WebApp render it

Cimba

Implemented by Andrei (deiu) Sambra in Q1

Provides basic microblogging features

Data stored as RDF (Turtle) files with LDP

Authentication done with WebID-TLS (for now)

Cimba Demo

Watch at: https://www.youtube.com/watch?v=z0_XaJ97rF0

Try it at: http://cimba.co

Play Chess via Blogs

1. Alice posts she's playing a game with Bob
2. Bob posts he's playing that game
3. Alice posts her move
4. (optionally) Bob posts a copy of that move
5. Bob posts his move … etc

- A client can render this like any chess app
- URLs linking all over the place
- Polling is slow; notification is good

Auction a textbook via Blogs

1. Alice posts her Chem textbook for auction
2. Alice notifies many auction index services
3. Bob finds it, posts a bid
4. Alice links to his bid as the new high bid

…

n. Alice closes the auction

- Again: UI can make this look like eBay, etc
- Note that eBay is more than software
- ISSUE: does Bob POST to his own blog and notify Alice, or does he POST to Alice's server?)

Machine Readable Blogs

These “POSTS” are RDF triples, not HTML

They are created and viewed through WebApps (or desktop or mobile apps)

They link to each other heavily, using Linked Data principles, much like Blog posts do.

Photo Sharing

1. Alice posts photos to her site
2. She tags, stars, groups, etc the photos
3. She sets ACLs on them
4. They get registered with index servers
5. Bob sees her photos in response to a search
6. He filters, tags, stars, etc some of them.

- They each see them in the UI they want, that they are using at the time.
- They have a default “skin” WebApp that makes them look nice to folks just following a link

Technologies in Cimba Today

Usual WebApp stuff

LDP (read and POST RDF (turtle) documents)

Vocabs: mostly sioc, some foaf

WebID (IRI for a person)
WebID-TLS (for now)

Technologies Cimba Still Needs

Who is following who? (backlinks)

- TrackBack, WebMention
- In SocWeb charter (currently in AC review)

Change Notification (avoid polling)

- revisit PubSubHubbub?

Query Service (follow many at once)

- Link-Following SPARQL

Still needs...

Better UX than WebID-TLS

Maybe a p2p CDN

Technology Summary

blogs

- people have their own sites
- their apps can post/edit their content
- backlinks
- change notification

+ semweb stack

- linked data platform (read/write RDF FS)
- link following sparql
- vocabulary conversion (owl?, rif?)
- indexing

+ identity = Crosscloud protocols

Is this just going to happen
on its own?

Probably not.

The Plan

Build an ecosystem

- users + apps + data servers, standards

Start small

- cimba, rww.io,

Viral growth

(easy, right?)

Perspective

1987 - Started to worry about this problem

1991 - My first startup in this space

2000 - EricP told me TimBL was on same track

2009 - RIF done, revisiting Social Lock-In

2010 - spare time=hacking on MVP

2012 - spare time=applying for grants

2013 - grants begin to arrive, WGs wrap up

2014 - assembling team

Grants

Very grateful to:

- John S. and James L. Knight Foundation
- QCRI
- National Science Foundation
- John D. and Catherine T. MacArthur Foundation

Research

Several Interesting Comp Sci Questions

- rdf patch
- link-following sparql
- efficient notifications
- scaling to many consumers (cf Akamai)
- scaling to many producers (cf Google)
- practical vocab conversion (RIF?)

Community

App Developers

- Start with MIT Students (UROP)
- Need to build a nice toolkit, docs, etc

Data Service Providers

- Partners? OwnCloud?
- Show it's a viable business

IndieWeb? FedSocWeb? tent.io?

Standards

Linked Data Platform (LDP)

- Going to CR now
- Crosscloud needs several other bits
- Hopefully re-charter in 6-12 mos
- … but they're not about social lock-in

SocWeb WG

- Hopefully launching in July
- … to soon to see their focus

Business

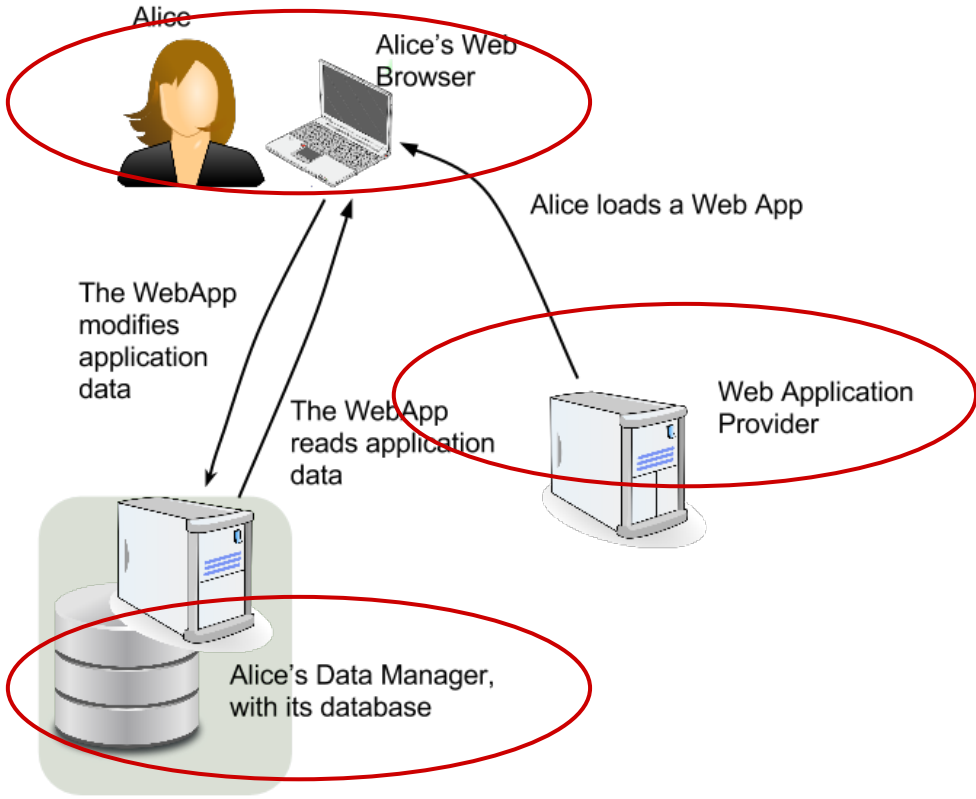
Ultimately, we need commercial participation

- In app development
  - but it won't be the folks that want to mine your data
  - it'll be the folks that want to provide value to users
- In cloud data hosting
  - intense market out there; this might catch on
- In data interoperability
  - apps written to be compatible with existing apps
  - apps written separately, integrated by another
- Branded and sold for end users
- Participating in standards

Customer Development

Multisided market, multiple customer bases:

1. End Users
2. Application Developers
3. Data Service Providers



End Users

Possible benefits:

- Sense of control, actual control
- Sense of privacy, actual privacy
- No ads (if you pay)
- Integrated interface to multiple systems
- Ability to include everyone
- Custom support for your group/enterprise

Issues:

- Need apps
- Need data service providers
- Need other users

Application Developers

Possible benefits:

- Doing right by the users
- Easier critical mass of users
- No need for backend operations
  - scaling, security, expertise, 24hr staff
- Ecosystem of collaboration

Issues:

- Need to hide RDF?
- WebApps only for now?
- Platform is still evolving rapidly

Data Service Providers

Possible benefits:

- Money
  - Different market from file storage
  - Not much harder
- Supporting your enterprise / people

Issues:

- Will be highly competitive market
- Platform is still evolving rapidly

Short Term

Focus on cimba MVP

- microblogging **plus**
- access control groups

and revise as needed

… While documenting protocols, feeding them into LDP and SocWeb WGs

… And doing necessary Comp Sci

Summary

- Social lock-in hobbles social software
- Crosscloud will solve social lock-in by
  - putting data under user control
  - in a standard format (RDF)
  - with lots of links
  - notification
  - indexes
  - vocab mapping shims
- Protocol/standards work needed
- Comp Sci research needed
- App developers needed

Next Steps

- Look at the app area lists and email me your favorites (and edits) https://github.com/sandhawke/crosscloud/blob/master/apps.md
- Join the announcements mailing list: crosscloud-announce@csail.mit.edu subscribe
- Let us know how your technologies/groups might fit in or use this
- Suggest more funding sources
- Suggest people to hire (especially postdoc, MIT UROP)