Confidentiality from the cloud provider

University of Birmingham Business Club
25 March 2014

Mark D. Ryan
University of Birmingham

Inexorable rise of cloud computing

Motivation for cloud computing

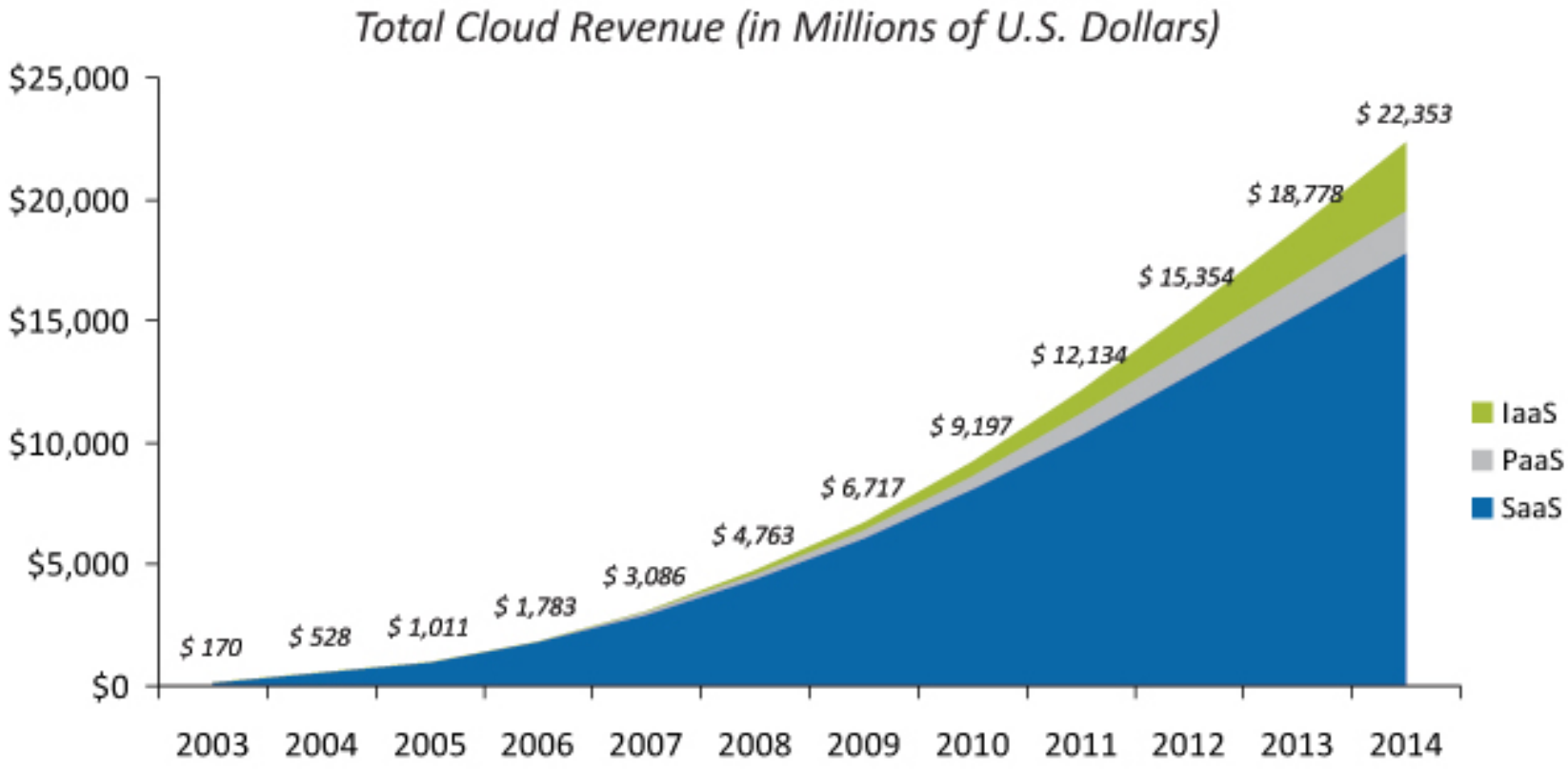
- simplicity
- cost
- security
- resilience
- flexibility
- pace of innovation

Worldwide cloud services revenue is forecast to reach $68.3 billion in 2010 and projected to reach $148.8 billion by 2014, according to a report by consultancy firm KPMG.

Consequences

- Google Apps is a packaged version of Google Docs, Mail, Calendar and other products under a custom domain name.
- In January 2011, 3,000 businesses were moving to Google Apps each day...
- ...and three million have moved since its debut in 2007.

Google Apps


Total Cloud Revenue (in Millions of U.S. Dollars)
$25,000
$20,000
$15,000
$10,000
$5,000
$0
$ 170
$ 528
$ 1,011
$ 1,783
$ 3,086
$ 4,763
$ 6,717
$ 9,197
$ 12,134
$ 15,354
$ 18,778
$ 22,353
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
IaaS
PaaS
SaaS

Eric Schmidt: Google Apps Has 40M Users; Adding 5K Companies Per Day

LEENA RAO

Thursday, September 1st, 2011 Comments

Google's Chairman and former CEO **Eric Schmidt** took the stage at CRM giant Salesforce's annual conference **Dreamforce** this evening. Salesforce founder and CEO Marc Benioff is interviewing Schmidt. Google's chairman started out with complimenting Salesforce how they've managed to become a company that defines modern enterprise computing, and that the company has the 'best vision for how enterprises will organize themselves."

Schmidt explained that enterprise customers can now be empowered with simpler solutions and don't have to access a complicated system, but thing are turnkey. "We went through a phase that is basic connectivity, then we have a connection and publishing phase (early parts of the web), and now we have a connecting phase," explains Schmidt.

Security: main obstacle to deployment

Availability and integrity - no problem

Google claims "no scheduled downtime" (thanks to distributed back end) and guarantees 99.9% availability.

Confidentiality - big problem

- Confidentiality violation by the cloud provider...
  - Google's business model involves mining its users' data.
- ...and its individual employees:
  - In July 2010, Google dismissed an engineer for accessing user accounts.
- ...and third-party adversaries
  - In December 2010, Google reported a "highly sophisticated and targetted attack" designed to steal information about users from Gmail.

The Cloud Security Alliance highlights loss of data confidentiality, malicious insiders, technology vulnerabilities and service hijacking as four of its seven top threats of cloud computing.

Security of cloud computing

Does user have to trust the service provider?

Security of cloud computing

Does user have to trust the service provider?

- Confidentiality ⟵ main issue
- Integrity
- Availability

NSA taps in to systems of Google, Facebook, Apple and others, secret files reveal

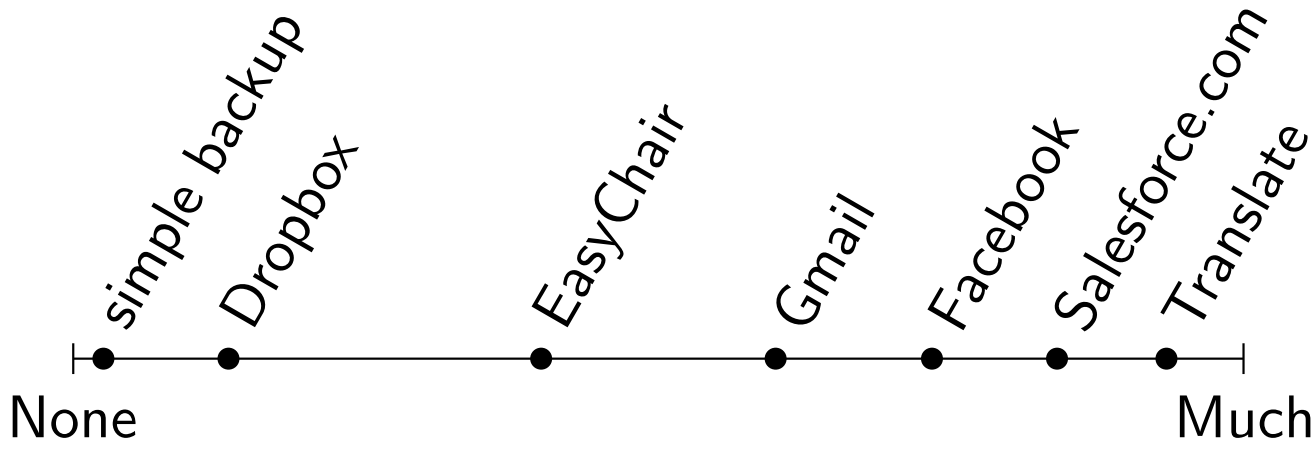
- Top secret PRISM program claims direct access to servers of firms including Google, Facebook and Apple
- Companies deny any knowledge of program in operation since 2007

Glenn Greenwald and **Ewen MacAskill**
The Guardian, Friday 7 June 2013

Jump to comments (1708)

Article history

Whether we have to trust the cloud provider or not depends on how much computation in the cloud is needed.



Avoiding having to trust cloud provider means:

- You don't have to trust its employees or subcontractors
- You don't lose data confidentiality if cloud provider gets hacked

Case study:

Trust domain over

Facebook

Case study:

Trust domain over

Case study:

Trust domain over

~~Facebook~~ **EasyChair**

Similar: **SalesForce.com**
FreeAgent
. . .

EasyChair: conference management

A **trust domain** consisting of authors, reviewers, & chairs

Year	#confs
2002	2
2003	3
2004	7
2005	66
2006	276
2007	629
2008	1312
2009	2183
2010	3305
2011	$\geq$ 4517

"We believe that since 2006 we have become number one conference management system in the number of conferences, users and submissions. All together EasyChair proudly hosted **29,000 conferences and 1M users**."

EasyChair: the confidentiality problem

- EasyChair managers have direct access to the submission and reviewing profiles of 1M users across 29k conferences, including submission rate, acceptance rate, reviewer profile (fair/unfair, thorough/scant, prompt/late).

EasyChair: the confidentiality problem

- EasyChair managers have direct access to the submission and reviewing profiles of 1M users across 29k conferences, including submission rate, acceptance rate, reviewer profile (fair/unfair, thorough/scant, prompt/late).

- EasyChair could [in principle, if it wished] offer profiling services to appointment panels, awarding bodies, recruitment agencies.

EasyChair: the confidentiality problem

- EasyChair managers have direct access to the submission and reviewing profiles of 1M users across 29k conferences, including submission rate, acceptance rate, reviewer profile (fair/unfair, thorough/scant, prompt/late).
- EasyChair could [in principle, if it wished] offer profiling services to appointment panels, awarding bodies, recruitment agencies.
- The data could become a target for hackers/crackers.

EasyChair: the confidentiality problem

- EasyChair managers have direct access to the submission and reviewing profiles of 1M users across 29k conferences, including submission rate, acceptance rate, reviewer profile (fair/unfair, thorough/scant, prompt/late).
- EasyChair could [in principle, if it wished] offer profiling services to appointment panels, awarding bodies, recruitment agencies.
- The data could become a target for hackers/crackers.
- Similar situation for other SaaS systems.

ConfiChair

ConfiChair: user experience

Aim: usability is as good

- Similarly to EasyChair, when you log into ConfiChair you should see all the conferences you are associated with.
- Users have accounts on ConfiChair; the names, affiliations, email addresses and other public data about users are known to ConfiChair.

Confidentiality from cloud provider

- User's browser encrypts data before sending it to ConfiChair, and decrypts data received from ConfiChair. Papers, reviews, discussions etc. are confidential from ConfiChair.

Prototype implementation

Implementation details

`confichair.org`

- The user interface follows that of EasyChair, HotCRP, and other conference management systems.
- Key generation, secure storage, protocol messages are all transparently performed by browser.
- Authors need to copy-paste $pub(Conf)$ from call for papers. Then, they just submit a paper using the usual "browse file system", "submit" buttons.
  - If an author submits more than one paper, the system remembers the key.
- Reviewers need to copy-paste K_{Conf} from their e-mail.
  - The system remebers the key after the first time it's used.

Implementation details

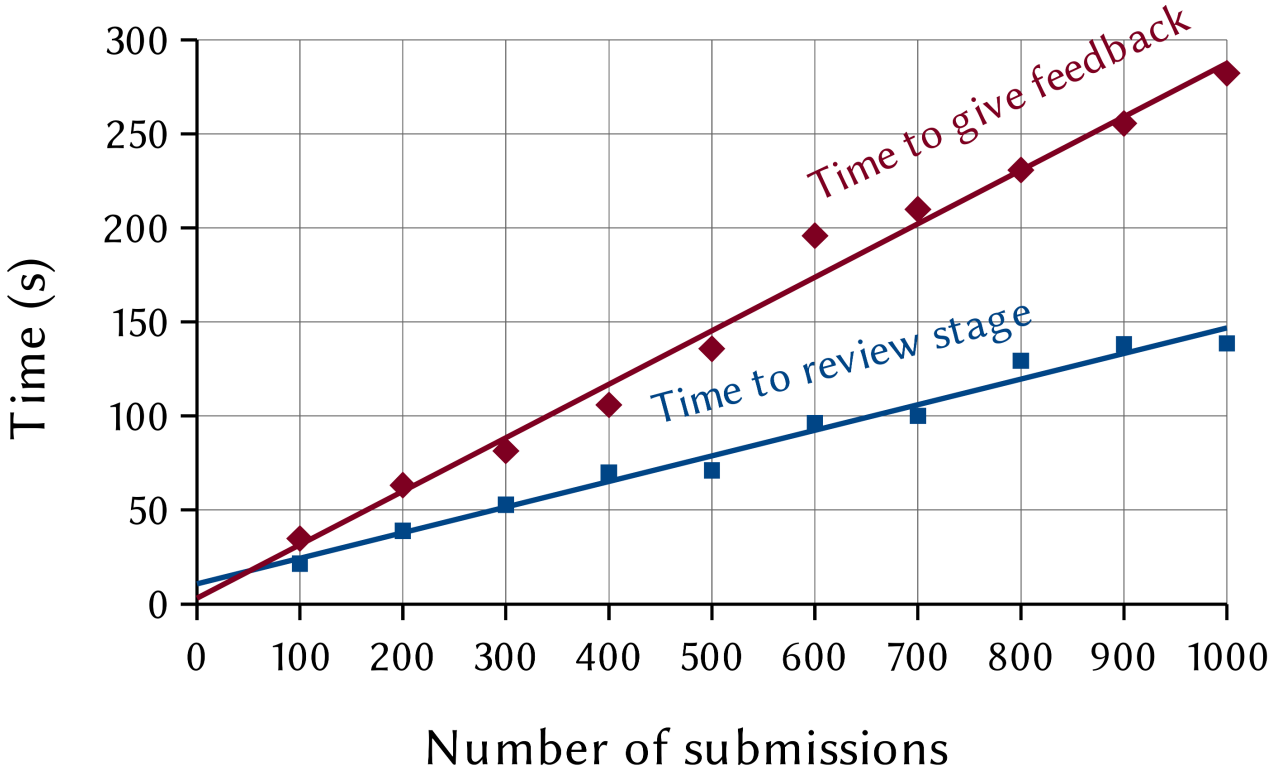
confichair.org

- The user interface follows that of EasyChair, HotCRP, and other conference management systems.
- Key generation, secure [illegible] all transp[illegible]
- A[illegible] pers. T[illegible] file sys[illegible]
- Revi[illegible] to copy-paste K_{Conf} from their e-mail.
  - The system remebers the key after the first time it's used.

Implementation:

- Uses HTML 5 features, incl local storage and W3C File API.
- Crypto by LiveConnect and Java plug-in.
- Crypto in Javascript, using SJCL.
- (Not yet done) Code signing by trusted party.

Key translation in the browser: performance



Speed evaluation. The time taken for moving to the review stage is about 75s for 500 papers. The time for moving to the feedback phase is about 150s for 500 papers.

Summary

ConfiChair

- ConfiChair is an infrastructure for hosting a type of **cloud-based trust domains**
- Its unique selling point is that the cloud doesn't see any sensitive data.
  - Therefore, it's invulnerable to attacks from/on the cloud provider, its employees, its subcontractors, or hackers that attack it.
- We formalised the properties, and verified them with ProVerif.
- Prototype implementation by Matt Roberts, Joshua Phillips, Mihai Ordean

`https://www.confichair.org/`

The future

- Production implementation
- Commercial exploitation

University of Birmingham Business Club

25th March 2014

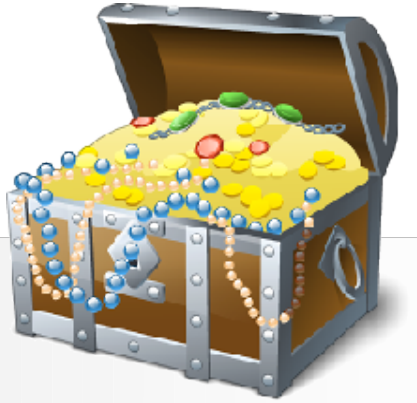
Total Data Care

Richard Zybert
Zybert Computing Ltd

What is Data Care

- Overview of data protection
- Partial solutions – what is available
- Personal view
- How do we do this

Do we care?



- We do.
- The general attitude to data security has changed dramatically in the last few years
- Most businesses rely on electronic data
- Current laws make it compulsory to care
- As time passes, we have all experienced data loss – we know the pain.

What is data security

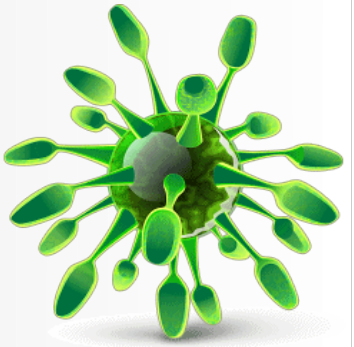
- Why do we protect anything?
  - Because we would miss it (data safety)

  - Because we don't want others to have it (data privacy)

- These are two separate issues and they apply differently to different kinds of data.

Protecting from data loss

- Organise
- Backup
- Stop viruses, trojans and other malware

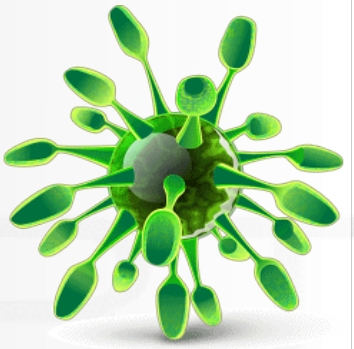


- Use software you trust

- Archive, Archive, Archive

Protecting data privacy

- Organise
- Control access
- Encrypt
- Stop viruses, trojans and other malware
- Use software you trust



Total Data Care

- Cover Files, Databases and Email
- Know where they are
- Protect from viruses
- Backup so you don't lose it
- Archive so that you can access old versions
- Control access so that you know who can read it

Planning and Control

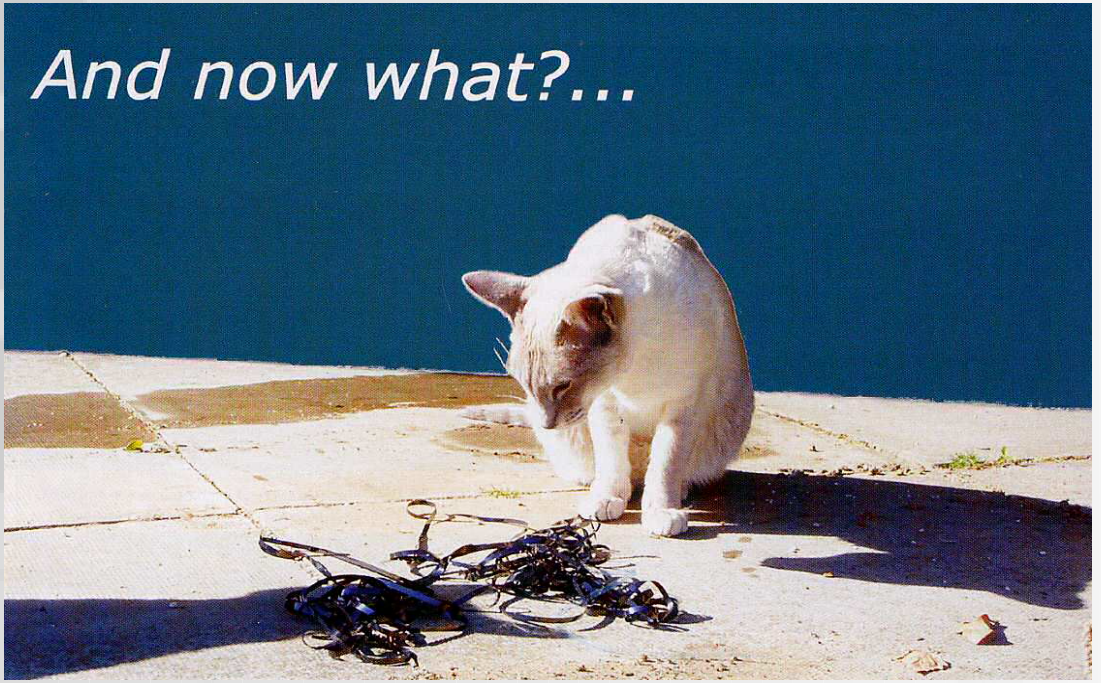
- This is the most complicated and time consuming part of the process
- Many businesses cannot afford to hire specialized consultants to help. It is usually expensive and rarely deals with the 'time consuming' bit
- Reducing the need for planning and control is the aim of many systems – some more successful than others

Enterprise solutions

- Many of the well established solutions have been originally designed long time ago
- Some of the design assumptions are not valid anymore
- 'Enterprise' usually means that
  - the system needs to scale to huge sizes
  - full time IT staff is available to manage it
- These assumptions lead to design, which may not be optimal for a smaller organization

Backup – what is on offer

- Magnetic tape
  - PRO - ... not sure ...
  - CON – slow, expensive, unreliable, time consuming



Backup cont...

- External disk (or NAS)
  - PRO
    - inexpensive
    - easy
  - CON
    - usually incomplete, no archiving
    - speed depends strongly on a solution
    - requires management
    - recovery may require special software
    - requires testing
  - Useful for ad-hoc, limited backup

Backup cont...

- On-Line backup
  - PRO
    - Inexpensive to start
    - No hardware to worry about
  - CON
    - Cost may rise with volume
    - Can get slow
    - Security not assured
    - Relies fully on service provider
    - Recovery slow and cumbersome

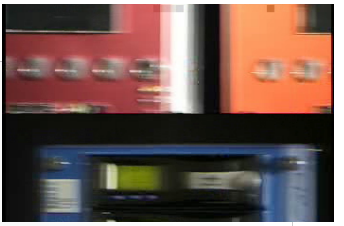
Backup cont...

- Document management systems
  - PRO
    - Full archiving and version control
    - Backup system usually included
    - Security can be controlled
  - CON
    - VERY expensive
    - Complicated
    - Requires professional planning and management

Why so much of IT is still insecure?

- Cost
- Complexity
- Inconvenience
- “No need”

So – what to do?



Zybert Computing approach is this:

- Use cloud for non-sensitive data and on-site for the rest
- Backup everything – so, no planning or control
- Archive as much as possible, including all email
- Provide secure remote access, so that nobody needs to carry data out of the office

This makes it possible for an organization to do its job without constantly worrying about the data