Impact Objective

- Explore the use of information flow control (IFC) to achieve greater security in cloud computing

Safety in the clouds

__Professor Jean Bacon__ is the __Principal Investigator__ of a project investigating end-to-end application security for cloud computing. Below, she explains more about the grant she and her team worked on, and how the project will help to improve data security

Can you begin by explaining what CloudSafetyNet is and how the Cambridge Flow Control Architecture (CamFlow) relates to it?

CloudSafetyNet: End-to-End Application Security in the Cloud is a grant awarded by the UK Engineering and Physical Sciences Research Council. CamFlow is software developed as part of the project. Some CamFlow software has been released as open source. This is an implementation of decentralised information flow control (IFC) within the operating system (OS), in the form of two Linux Security Modules: one to enforce IFC, the other to perform auditing of attempted data flows – both those allowed and those forbidden by IFC. CamFlow therefore enforces and audits data flows at the OS level, where data flows between software running on the same OS.

We have also extended our message-passing middleware (SBUS) to incorporate IFC. Middleware is used to pass data between software running on different OSs – between software running on different machines within a cloud and for external communication. This includes interactions with end users and flows of data from devices and sensors into a cloud, and potentially the Internet of Things (IoT) where cloud services are a component part. However, the OS run by the cloud is under the control of the cloud provider, so it can be trusted to enforce IFC, whereas IoT OSs cannot be trusted to enforce IFC reliably. We therefore consider this branch of the project as future research, although we have published ideas on the challenges involved and how to tackle them.

Why are current cloud data security measures insufficient?

The original designers of clouds saw a major reason for lack of uptake as a data protection problem: mutually suspicious tenants having to share the same cloud – computers and data stores. Not only might a tenant's software be running on an adjacent machine to another tenant's, it might be sharing the same machine. For this reason, cloud designs offered 'strong isolation' of tenants, enforced by 'virtualisation' – each tenant's software (including an OS) runs in a separate 'virtual machine' (VM) above a layer of software called a hypervisor that runs on the hardware. This ensures that the running software of one tenant is completely separated from that of another. This style of cloud use is called Infrastructure as a Service (IaaS). To avoid having many copies of OSs running, one in each VM, cloud providers offer Platform as a Service (PaaS), where separate applications run in isolated containers above a common, cloud-provided OS. Cloud architecture is still an active research area – library operating systems or unikernels run directly over the hypervisor and ensure that each tenant's OS contains only the functions needed by that tenant. However, data may need to flow between different cloud tenants, such as different applications running on behalf of the same end user, so strict isolation is often too rigid. What is needed is strictly controlled sharing of data between tenants, as specified in end users' policies. IFC achieves both protection and controlled sharing of data.

How are you communicating with your stakeholders and the wider public about the results from this project?

We have published the ideas in international conferences and journals. A 'Big Ideas' paper was presented, by CloudSafetyNet Research Associate Jatinder Singh, at the ACM International Middleware conference in December 2016, which emphasised the impact of the work and new research directions. While the OS aspect is available as open source, the CamFlow middleware needs to be fully integrated with IFC before a working system can be supported. Student projects are being set up to use CamFlow both in Cambridge and in Harvard, where another CloudSafetyNet researcher, Thomas Pasquier, has taken a postdoctoral research position. His current interest there is to establish data provenance via IFC audit.

Cloud data management poses a huge problem for international law. Can a system such as CloudSafetyNet be used to address this issue?

We have experimented with how to use IFC to enforce regulations about location, such as 'personal data created in the EU must not leave the EU'. We can see how to enforce this using simple IFC tags on data and authorised locations. This is an important area. We have collaborated – as part of a different project funded by Microsoft – with Queen Mary University of London's Department of Commercial Law to launch the virtual Microsoft Cloud Computing Research Centre. We have also set up an annual workshop, as part of the IEEE International Conference on Cloud Engineering. The CLaw (Cloud Law) workshop will run for the third time in April 2017.

Securing and controlling information flow in the cloud

*The **CloudSafetyNet** project focuses on fundamentally rethinking how cloud computing can handle the security requirements of applications. By providing software that audits and checks the flow of data, the team hope to protect against security violations and vulnerabilities in the cloud itself*

Cloud computing is a practice whereby users rely on a network of remote servers hosted on the Internet to store, manage and process data, as opposed to using local servers or personal computers. There are many benefits to cloud computing: it enables companies to focus on their core business, instead of spending time and money on computing infrastructure; cloud resources can expand to meet peak demands and contract at quiet times (so-called elasticity); and the lack of maintenance associated with cloud computing means companies are often able to get their business established quicker and more cheaply.

Despite the obvious benefits of cloud computing, it is not without its problems. That the service provider is able to access the data it stores in its cloud at any time raises obvious privacy concerns, but could also – theoretically at least – lead to the accidental (or deliberate) deletion of information, or retaining data that should have been deleted. Because cloud infrastructures combine services and software written by a variety of people and teams, there is no unified approach to guaranteeing data security. Despite these potential pitfalls, cloud computing represents a revolution in the ways in which companies, research institutions and government organisations can offer applications and services to users in the digital economy. For this reason, many researchers are looking at ways to ensure the security of cloud computing, so that it can realise its extremely promising potential.

ALLAYING FEARS OVER DATA BREACHES BY CONTROLLING FLOW

With that in mind, the UK Engineering and Physical Sciences Research Council awarded a grant to a team of researchers in 2013. The CloudSafetyNet: End-to-End Application Security in the Cloud project involves researchers from the University of Cambridge and Imperial College London, with collaborators from the Cancer Registry, Public Health England and Otago University in New Zealand. After the Cambridge grant ended in mid-2016, one researcher moved to a postdoctoral position in Harvard, where the research continues. Principal Investigator Professor Jean Bacon hopes to create end-to-end application security in the cloud through the use of decentralised information flow control (IFC). IFC is a data-centric security mechanism that can enforce and track information flow to improve security in the cloud in a variety of ways. For instance, it enables the cloud platform to impose checks that enforce security policies and track data flows across different services, thereby improving accountability of data, in terms of where it was created and what has happened to it – its provenance. The team behind CloudSafetyNet has already shown how IFC can be deployed as part of a cloud-provided operating system (OS).

RELATIONSHIP WITH CAMFLOW

The team has designed IFC in such a way as to ensure that any application need not necessarily be aware of it; that the application development process need not be more complex; and that legacy software can run unchanged. However, tenants and their end users can choose to specify policy on how their data can flow.

While cloud service providers are usually large, well-known companies – and are therefore likely to be trusted by end users – single cloud tenants that provide services for end users do not have that benefit. However, with cloud-OS-provided IFC, applications do not need to be trusted in this way because they are compelled to use the OS-enforced

"We have experimented with how to use IFC to enforce regulations about location, such as "personal data created in the EU must not leave the EU". We can see how to enforce this using simple IFC tags on data and authorised locations"

IFC mechanism. In short, the team's design of IFC has cured many potential problems by enforcing compliance with security mechanisms.

One of the most exciting outputs from the project has been the team's creation of the Cambridge Flow Control Architecture (CamFlow), which is a pair of Linux Security Modules (LSMs) – one for IFC and the other for audit data capture.

IFC AS A SAFETY NET FOR DATA

While IFC's main focus is not on security attacks, it can be a useful addition to other security technologies in detecting, confining and analysing – through audits – some types of security attack. For example, IFC can ensure that any input code and data must go through a validation process before flowing further into a system, which could obviously prevent some specific problems. However, IFC should be seen as a means of confining the flow of data, and ensuring its quality and authenticity, as opposed to countering network-based attacks such as distributed denial of service (DDoS). The ultimate focus of IFC is on data leaks, where tagged data can only flow to similarly tagged entities, thereby enabling the audit of all such flows. Through audit, the data can be tracked to see if it has been leaked and, if it is claimed there has been a leak, IFC audit can either provide evidence on where the data flowed to support the claim, or dismiss it.

This is all achieved using two LSMs – one for IFC enforcement and one for IFC audit – that have been added to Linux OSs used in clouds. These check every call to the OS that involves the transfer of data, enabling the team to achieve a form of mandatory access control (MAC). 'There are two types of IFC tag – those for privacy and confidentiality, and those for integrity, quality and authenticity,' explains Bacon. 'The LSM checks that the destination of the attempted data flow has secrecy tags that indicate it is equally or more trusted for secrecy than the source, and that the integrity tags indicate that the source is equally or more trusted for integrity than the destination.'

BEYOND A SINGLE OS – MIDDLEWARE

For data flows between software running on different OSs – within or external to a cloud – middleware is needed. The team have extended their messaging middleware to incorporate IFC, but further work is needed before this is released as open source software. Such middleware supports interactions within the Internet of Things (IoT), that is, the internetworking of physical devices and other items that are embedded within such things as electronics, software and sensors to enable the collection and exchange of data. 'Clouds may be used as components of the IoT – for example to hold data gathered from sensors or to offload processing from low-powered devices,' explains Bacon. '

While there is recognition that the strands of the project relating to IoT are beyond the scope of CloudSafetyNet, the team has made significant headway in providing software that acts as a safety net to protect against security violations caused by flaws in applications or vulnerabilities in the cloud itself. While realising such potential is by no means simple, CloudSafetyNet represents a significant step in the right direction towards a secure future.

Project Insights

FUNDING

Engineering and Physical Sciences Research Council (EPSRC), UK

CONTACT

Jean Bacon
Principal Investigator

T: +44 122 3334604
E: jmb25@cl.cam.ac.uk
W: http://www.cl.cam.ac.uk/research/srg/opera/projects/csn/

PRINCIPAL INVESTIGATOR BIO

Professor Jean Bacon is Professor Emerita of Distributed Systems at the University of Cambridge. She was the first woman to be appointed as a Lecturer in the Computer Laboratory in 1985. Bacon led the Opera research group, focusing on open, large-scale, secure, widely distributed systems. Collaborators include the Cancer Registry, Public Health England and the Microsoft Cloud Computing Research Centre. She is a Fellow of Jesus College, of the IEEE and the BCS, and was a Governing Body member of the IEEE Computer Society from 2000 to 2007. She is on the steering committee of a number of conferences including the IEEE International Conference on Cloud Engineering. She is the author of *Concurrent Systems: An Integrated Approach to Operating Systems, Distributed Systems and Databases* and holds an Honorary Doctorate from the Open University.

UNIVERSITY OF CAMBRIDGE
Computer Laboratory

EPSRC
Engineering and Physical Sciences Research Council