

出國報告（出國類別：國際會議）

## 參加 2015 年雲端運算國際研討會 出國報告

服務機關：行政院主計總處主計資訊處

姓名職稱：余分析師威逸

派赴國家：美國

出國期間：104 年 6 月 25 日至 7 月 3 日

報告日期：104 年 9 月

# 摘 要

電機電子工程師學會（Institute of Electrical and Electronics Engineers，簡稱 IEEE）是國際著名之專業技術組織，其主辦之雲端運算研討會是各國學者與專家交流研究成果與實務經驗之重要論壇，今年以「服務」為主題，研討如何運用雲端運算技術發展 XaaS（everything as a service）的雲端服務環境，並針對雲端服務品質、安全及隱私、移轉、節能及相關應用於雲端服務之技術研究深入探討。

根據 Forrester 研究，全球的雲端運算市場將從 2011 年的 40.7 兆美元成長到 2020 年的 241 兆美元[2]。此外 Gartner 也預測公有雲端服務花費從 2011 年至 2016 年，預估每年將有 17.7%成長率[3]，台灣雲端運算服務市場也在政府推波助瀾下迅速成長。而本總處近年來為持續關注雲端運算服務技術最新發展趨勢並提升相關應用規劃能力，業於 104 年度預算提報出國計畫在案，擬派員參加 IEEE CLOUD 2015 國際研討會後，將最新發展雲端運算服務的技術分享本總處同仁，提供後續雲端服務規劃參考。本報告內容計分四章，第一、二章為參加本次研討會之目的與過程，第三章則是研討會重要內容摘要，第四章為提出本次參加研討會的心得與建議。

# 目 次

一、目的 .....	1
二、過程 .....	2
三、會議重要內容摘要.....	6
四、心得與建議 .....	14
參考文獻 .....	16

## 一、目的

以往傳統資料中心具有複雜實體架構、高營運成本、低彈性及不容易擴充等缺點，由於近年來雲端運算服務崛起，其虛擬化技術允許多台作業系統建置於一台實體機上，除可解決傳統資料中心上述問題外，又具有高可靠度及經濟實惠優勢，由於這些優勢因素，促使許多私人企業因上述優勢及成本考量，慢慢將企業內部的資料中心及應用程式移轉至雲端運算服務，而這股趨勢也趨使行政院花費數十億元打造 10 朵政府雲，本總處近年來積極投入建置主計雲，提供主計人員使用共通性歲計會計系統，總而言之雲端運算服務已成為未來資訊科技共同發展趨勢。

雲端運算服務技術除能提升軟硬體資源共享效益及資訊服務之便利性，也與本總處近年來歲計會計系統移轉至雲端作法息息相關，為持續關注雲端運算服務技術最新發展趨勢並提升相關應用規劃能力，業於 104 年度預算提報出國計畫在案，擬派員參加由電機電子工程師學會（Institute of Electrical and Electronics Engineers，簡稱 IEEE）主辦之 IEEE CLOUD 2015 國際研討會，將最新雲端運算服務所使用的技術、規劃雲端運算服務所需考量要素，提供本總處及政府單位後續規劃雲端運算服務參考。

## 二、過程

本次研討會（IEEE Cloud 2015）與 ICWS（International Conference on Web Service）、SCC（International on Services Computing）、BigData Congress（International Congress on Big Data）、MS（International Conference on Mobile Services）及 SERVICES（World Congress on Services），針對不同主題共計舉辦 6 個主要國際論文研討會，同時於 104 年 6 月 27 日至 7 月 2 日在美國紐約市 Millennium Broadway Hotel 同時舉行，分不同會議室舉辦，本次研討會會議議程如表 1 及部份照片集錦如圖 1。

表 1：研討會會議議程

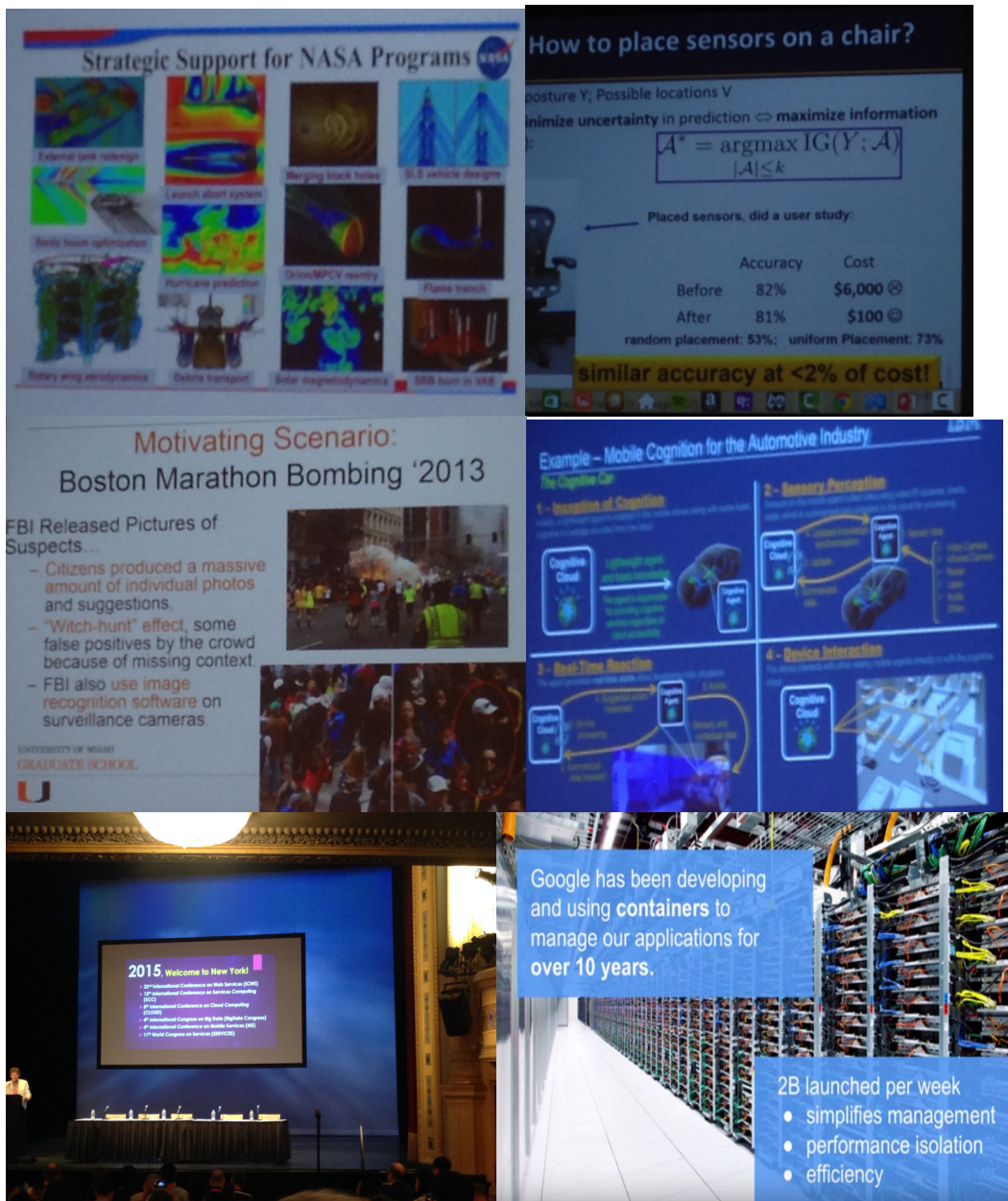
時間\日期	6月27日	時間\日期	6月28日	時間\日期	6月29日
09:00~10:00	CLOUD Short Paper 1	08:15~09:15	CLOUD Research 1	08:15~09:15	CLOUD Research 3
10:00~10:15	Break	09:15~09:40	Break	09:15~09:25	Break
10:15~11:15	CLOUD short Paper 2	09:40~11:30	A Holistic View of Software Evolution toward Software as a Service, Carlo Ghezzi	09:25~10:25	CLOUD Visionary Track2
11:15~11:20	Break	11:30~12:30	Lunch(not included)	10:25~10:50	Break
11:20~12:20	CLOUD Short Paper 3	12:30~13:30	CLOUD Research 2	10:50~12:00	Big Data as Service at NASA, Tsengdar J. Lee
12:20~13:30	Lunch(not included)	13:30~13:30	Break	12:00~13:00	Lunch(not included)
13:30~14:30	CLOUD Short Paper 4	13:30~14:40	Services Computing(Moderators: John Miller, Hong Zhu	13:00~14:00	CLOUD Research 4
14:30~14:40	Break	14:40~14:50	Break	14:00~14:10	Break
14:40~15:40	CLOUD Short Paper 5	14:50~15:50	CLOUD Visionary Track 1	14:10~15:20	Service Economics on CLOUD Computing(Moderator: Ajay Mohindra)
15:40~16:10	Break	15:50~16:15	Break	15:20~15:30	Break
16:10~17:10	CLOUD Short Paper 6	16:15~17:25	Big Data and IoT(Moderator: Tony Shan)	15:30~16:30	CLOUD Research 5

表 1 : 研討會會議議程

時間\日期	6 月 30 日	時間\日期	7 月 1 日	時間\日期	7 月 2 日
08:15~9:15	CLOUD Research 6	08:15~9:15	CLOUD Research 10	08:15~09:15	CLOUD Research 14
09:15~09:25	Break	09:15~09:25	Break	09:15~09:45	Break
09:25~10:25	CLOUD Research 7	09:25~10:25	CLOUD Research 11	09:45~10:45	CLOUD Research 15
10:25~10:50	Break	10:25~10:50	Break	10:45~10:55	Break
10:50~12:00	Data Analytics as a Service: The Next Big Challenge in Service, Ling Liu	10:50~12:00	Cognitive Computing at IBM Research, Guruduth Banavar	10:55~11:55	CLOUD Research 15
12:00~13:00	Lunch(not included)	12:00~13:00	Lunch(not included)	11:55~13:00	Lunch(not included)
13:00~14:00	CLOUD Research 8	13:00~14:00	CLOUD Research 12	13:00~14:00	CLOUD Research 16
14:00~14:10	Break	14:00~14:10	Break	14:00~14:10	Break
14:10~15:20	Kubernetes and the Path to CLOUD Native, Eric Brewer	14:10~15:20	The CLOUD Software Stack: From SDN to Micro-Services, Dennis Gannon	14:10~15:00	Closing Session
15:20~15:30	Break	15:20~15:30	Break		
15:30~16:30	CLOUD Research 9	15:30~16:30	CLOUD Research 13		
16:30~16:50	Break	16:30~16:50	Break		
16:50~18:00	Convergence of CLOUD Computing and Big Data : Making Big Social and Human Impact(Moderators: Calton Pu)	16:50~18:00	Mobile and IoT Services(Moderator: Nimish Radia)		

資料來源：IEEE CLOUD 2015

圖 1：研討會會議照片集錦



資料來源：IEEE cloud 2015



### 三、會議重要內容摘要

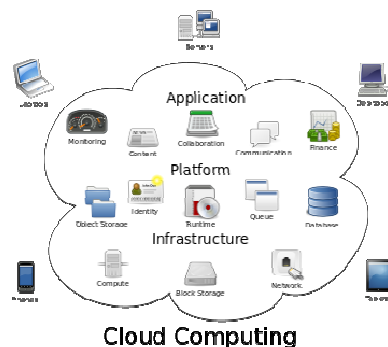
本次研討會議主要內容以邀集論文發表 (Paper Presentation) 及知名專家演講 (Opening Plenary Session) 二種形式進行，以雲端運算服務研究論文發表為主，專家演講為輔，專家演講部分，則邀請業界重量級人物參與演講，如：Google 負責基礎設施的副總裁 Eric Brewer、IBM 認知運算研究部副總裁 Guruduth Banavar 等。本次研討會論文發表部分，以雲端服務為主軸，探討雲端服務模式、服務品質、安全及隱私、移轉、節能及雲端服務技術等雲端領域相關議題；專家演講部分，包含雲端運算服務及大數據資訊科技應用等議題。經歸類綜整後，將針對本次研討會雲端運算服務五大重點予以說明。

#### 3.1 雲端運算定義、服務模式及實例

##### 1. 雲端運算定義

雲端運算 (Cloud Computing, 如圖 2)，是一種基於網際網路的運算方式，透過這種方式，使用者可以透過共享軟硬體資源來降低成本，達到經濟效益。大致可以分兩大類，一為「雲端技術」，即利用虛擬化、管理自動化 (如：版本佈署等)、分散式儲存及處理、加密等技術，來實現雲端服務目的，另一為「雲端服務」，即透過網路所提供服務模式，可分為軟體即服務、平台即服務及基礎設施即服務三項。

圖 2：雲端運算架構圖



資料來源：wikipedia[1]

## 2. 雲端服務模式

目前雲端服務佈署類型依美國國家標準和技術研究院可分為：公有雲、私有雲、社群雲、混合雲四種型，其定義可參考[1]，而服務模式共區分為三種類型，茲說明如下：

### (1) 軟體即服務 (Software as a Service, 簡稱 SaaS)

軟體即服務有時稱為「即需即用軟體」(即是一經使用者要求，即可立即使用)，它是一種軟體交付模式[1]。這種服務是在雲端安裝軟體後，使用者只需透過網際網路，不用安裝軟體即可使用。而用戶通常使用客戶端經由一個精簡網頁瀏覽器來使用軟體。如：Facebook、Youtube、Gmail、DaaS (如：iCloud 提供雲端硬碟服務均是屬於 Data as a Service) 均是屬於此類型服務。

### (2) 平台即服務 (Platform as a Service, 簡稱 PaaS)

過去公司企業需購置昂貴的程式開發軟體，如 MS .Net Studio，而 PaaS 介

於軟體即服務與基礎設施即服務之間，提供使用者另一選擇，提供使用者於開發軟體時所需環境及工具，使用者無須煩惱其作業系統、網路、防毒軟體更新等，專注於程式開發，簡化使用者作業流程。如中華電信推出之 PaaS 服務，提供 Java、.NET 及 PHP 三種運行環境。

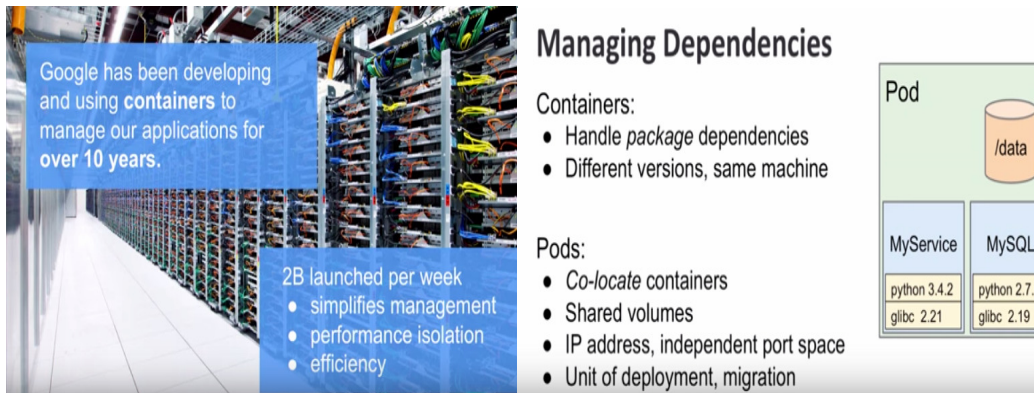
### (3) 基礎設施即服務 (Infrastructure as a Service, 簡稱 IaaS)

IaaS 不像傳統模式一般，使用者需自行建置 IT (Information Technology) 基礎設施，如：機房、防火牆、伺服器，花費昂貴建置成本，IaaS 即提供使用者另一選擇，使用者無須購置軟體、硬體等網路設備，即可任意部署網路元件、快速建置作業系統，並可以控制作業系統、磁碟及佈署其應用程式等。如：Amazon Elastic Compute Cloud (EC2) [4]及 Windows Azure [14]、Rackspace [15]及 Google Compute Engine [11]等，可以自由選擇其所需作業系統、及所需搭配的各项應用軟體及資料庫。

## 3. 雲端服務實例

Google 負責基礎設施的副總裁 Eric Brewer 敘述 Google 已經發展並使用 Linux Containers 來管理其應用系統 (如：Google Mail 或 Map 等) 超過 10 年了，使用 Linux Containers 的原因，則是具有效率及效能的優勢。另程式語言使用 Python (語法類似 C 程式語言)、關連式資料庫使用 MYSQL 資料庫，其軟體、作業系統於標準化後更容易管理，Google 每個禮拜佈署其應用程式於 20 億台虛擬主機上 (Google 機房及系統架構如圖 3 所示)。

圖 3：Google 機房及系統架構圖



資料來源：Eric Brewer 投影片

### 3.2 雲端服務品質

2014 Cloud service outage (雲端服務中斷) [9][10]統計 Amazon EC2 去年共遭遇到 20 次服務中斷，計 2.4 個小時；微軟 Azure 去年共遭遇到 92 次服務中斷，計 10.97 個小時；Google GCE 去年共遭遇到 72 次服務中斷，計 4.42 個小時，而其雲端服務的品質，有賴於雲端運算服務供應商與使用者於服務水平協議 (Service Level Agreement；SLA) 規範，而服務水平協議即規範提供雲端服務廠商於雲端運算服務使用者所需保證服務品質 (Quality of Service，QoS) 所達成的協議，其協議可以是合約一部分，其 QoS 可以要求所應保留的網路頻寬、可靠度、延遲、平均修復完成、無故障時間等。

### 3.3 雲端服務虛擬化軟體比較

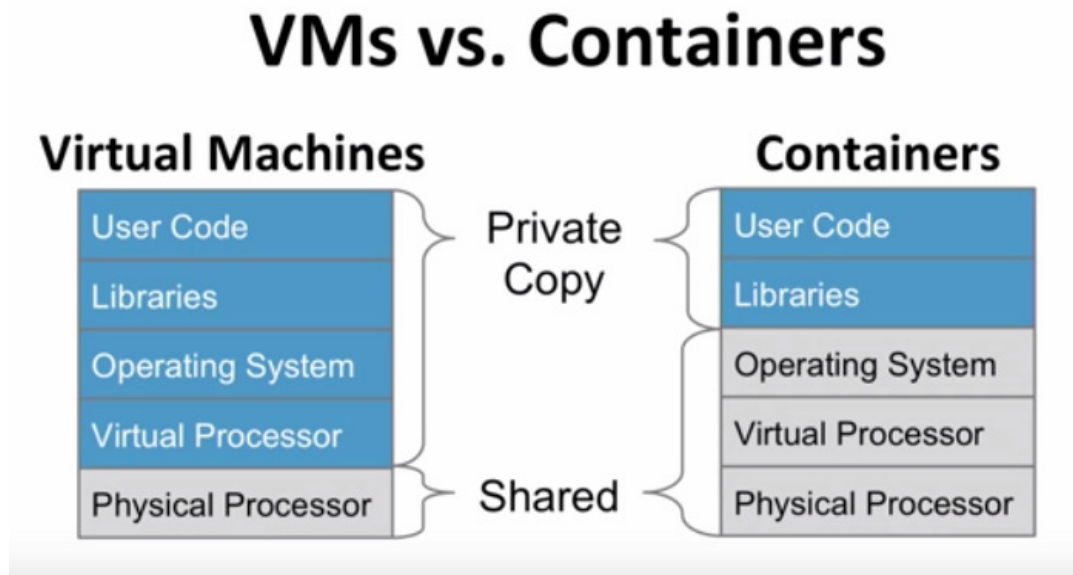
所謂虛擬化技術便是允許一台實體伺服器上架設多台虛擬作業系統技術，藉由共享資源來降低成本，而目前雲端運算服務所使用虛擬化技術分為 VMs (Virtual

Machines)及 Containers 均是雲端運算服務所使用的虛擬化技術。

近期研究報導[16]針對 Dock Containers 與 VMs 進行比較，Docker Containers 只消耗約 6%及 16%的 VMs 之 CPU 及記憶體使用率，且它開機時間及重開機時間約佔 60%及 2%的 VMs 之反應時間。故相較 VMs，由於 Containers 使用較少的記憶體(VMs 和 Containers 記憶體配置比較如圖 4)及 CPU，故可以執行較多的應用程式，也具有較佳執行效能，而 Containers 的缺點在於安全性，由於在 Containers 上的應用程式具有最大權限，不當的應用程式極可能將作業系統癱瘓，核心作業系統也易被後門程式入侵。

選擇 VMs 或 Containers 之時機，茲說明如下：如何決擇到底要使用 Containers 或是 VMs，如果應用程式像是 Google 一樣，如：單一應用程式安裝及佈署在數量較多台虛擬主機上，則 Containers 是較佳的解決方案；相反地如需佈署多個不同應用程式（如：IIS 及 Apache 等），則 VMs 具有彈性優勢，故 VMs 則是較佳選擇。

圖 4：VMs 和 Containers 記憶體配置比較圖



資料來源：Eric Brewer 投影片

### 3.4 其他雲端服務重要議題

#### 1. 安全及隱私

一般企業將資料中心移至雲端服務時，為了保護個資及機敏性等資料，在資料移至雲端服務前，會將資料加密，來確保其個資及機敏性等資料不外洩。而其加密解密的技術不外乎公、私鑰技術，或其混合技術，其目的除保護個資及機敏性等資料，也為了防止駭客入侵竊取資料。

#### 2. 移轉(Migration)

虛擬主機移轉即是由一台實體節點，將其主機上的虛擬主機移至另一台實體節點上。可分為線上及離線移轉，線上即不影響其所提供的雲端服務，但會產生安全及佔有頻寬資源問題；而離線移轉，其原雲端服務會中斷，其作法是先暫停來源端主機雲端服務，待來源端的所有虛擬主機移至另一台實體節點上後，再由目的端主機提供雲端服務。

#### 3. 節能

全美資料中心在 2013 年共計使用了 910 億度電，預計 2020 年將使用到 1,400 億度電[19]，高電費不僅產生高營運成本，也造成大量的二氧化碳被排出。如何節省電的花費，其方法有低耗能晶片（如：dark-silicon）的使用、動態關機（Dynamic Shutdown，DNS）、透過資源管理平台來有效地運用資源以節省能源。由於 Containers 相較 VMs 使用較少的 CPU 及記憶體，使用 Containers 虛擬化技術較使用 VMs 更能有效節省能源，也由於此原因，VMs 的虛擬化技術也將 Containers 包裝在 VMs 內部，由此可知 Containers 虛擬化技術也日趨重要。

### 3.5 雲端技術常見開放原始碼軟體

開放標準對於雲端計算的發展佔有極重要角色，且開放原始碼已經為眾多的雲端計算實例提供基礎架構，茲針對開放常用雲端運算的軟體說明如下：

#### 1. Docker

Docker [17]是一個開放原始碼軟體，讓應用程式自動佈署在Linux Containers下，在Linux作業系統上提供一個額外的軟體抽象層，以執行作業系統層虛擬化的自動管理機制。

#### 2. OpenStack

OpenStack [20]是一個美國太空總署和Rackspace合作研發的雲端運算軟體，是一個由處理、儲存、網路等不同功能的內部元件所組成開放原始碼軟體。它是IaaS（基礎設施即服務）軟體，讓任何人都可以自行建立和提供雲端運算服務。

#### 3. Hadoop

Hadoop [8]為利用分散式儲存及運算(MapReduce)來處理大量資料集(BIG DATA)所開發出來開放原始碼軟體框架，其中核心元件所組成：一為MapReduce [18]，另一為Hadoop Distributed File System (HDFS) [12]。MapReduce[18]：應用程序被分割成許多小部分，而每個部分都能在群體中的任意節點（如：虛擬主機）上執行。HDFS：是一個分散式、可擴充、可攜式的檔案系統，藉由在不同節點重製資料以提供資料可靠度。其它和MapReduce相似軟體，如Storm [6]、Spark [7]等。

由於 NOSQL 具有高可靠度、彈性及擴充性。HBase [5] 是一個 NOSQL 應用實例，它參考了 Google 的 BigTable 建置，屬 Hadoop 的一部分，使用在 HDFS 文件系統上，而臉書便是使用 HBase 來管理大量數據資料

#### 4. 其他雲端軟體

管理雲端上千萬台虛擬主機是一大問題，而問題的解決方法仰賴其自動化佈署工具，常用雲端自動化佈署工具有：Puppet [23]、Chef [22] 及 Ansible [21] 等，它提供新增使用者、安裝程式、設定 IP 等功能，可以在短時間完成多台虛擬主機的系統設定及程式安裝等。以 Puppet 為例，其運作時每台虛擬主機都需安裝 Puppet Agent，每一台的 Puppet Agent 都需向主控台伺服器回報現行應用系統的運行現況、佈署情況等訊息，管理者可即時了解系統運行及佈署程式狀況。



## 四、心得與建議

本次參與 2015 年雲端運算國際研討會，除可瞭解現行雲端服務發展趨勢並提高資訊人員視野外，對於未來規劃建置雲端服務之資訊人員能夠獲得不少啟發。以下就從本次研討會內容所啟發之心得，提出規劃資訊發展策略一些看法以供參酌。

### 1. 標準化開發與維運環境，優化雲端服務效能。

為利主計雲端服務佈署及維運，本總處應規劃主計應用服務標準化開發規範，讓集中維運環境一致，並善用自動化管理軟體（如：Puppet 等），以期各服務可迅速佈署至雲端。標準化環境除可擷節主計資訊系統開發及維運成本外，亦可使各服務易於整合再利用，以達成資訊資源共享目標。

### 2. 落實定期演練機制，提升主計服務品質。

為確保資料於災害或意外後能順利回復，本總處須落實定期演練機制（如：線上及離線移轉虛擬伺服器演練），確保主計人員雲端服務品質。本總處目前提供離線移轉演練，建議演練時納入線上即時移轉虛擬伺服器演練，提供不間斷雲端服務。另有關不間斷雲端服務衡量品質，則應採其國際上衡量品質標準，並善用自動化工具統計一年內服務中斷次數及合計小時，並與國際上 Amazon EC2 及微軟 Azure 相比，確保主計人員享有國際級的雲端服務。

### 3. 強化資安防護機制，確保雲端服務安全。

資訊安全雜誌[13]報導駭客於 Amazon EC2 的 Linux 平台上放置後門程式 Distributed Denial of Service(DDOS) Trojan-Backdoor.Linux.May.g 後，

發動 DDOS 攻擊。雲端運算服務採用虛擬化技術與架構，所衍生的資安管理與攻擊防禦機制皆比以往傳統資安領域更為複雜。本總處除現行採用之防火牆和入侵偵測/防禦系統等安全防護機制外，應考量運用自動化軟體來監督及管理虛擬伺服器。另為確保雲端資料安全，建議於資料庫移至雲端服務前，須將密碼、個資或機敏性資料予以加密處理，且需使用多重安全認證機制（如：自然人憑證等）登入及管理虛擬伺服器，增加資訊安全強度。

## 參考文獻

- [1] [https://en.wikipedia.org/wiki/Cloud\\_computing](https://en.wikipedia.org/wiki/Cloud_computing).
- [2] <http://www.zdnet.com/blog/btl/cloud-computing-market-241-billion-in-2020/47702>.
- [3] <http://www.datacenterdynamics.com/focus/archive/2014/05/empowering-service-providers-engineer-cloud-intelligent-networks>.
- [4] Amazon. Amazon Elastic Compute Cloud (Amazon EC2). <http://aws.amazon.com/ec2>
- [5] Apache HBase website. <http://hbase.apache.org/>.
- [6] Apache, “Storm: Distributed and fault-tolerant realtime computation.”[Online]. Available: <https://storm.incubator.apache.org/>
- [7] Apache, “Spark: Lightning-fast cluster computing.” [Online]. Available:<https://spark.apache.org/>
- [8] “Apache Hadoop,” <http://hadoop.apache.org>.
- [9] Brandon Butler, “Cloud providers are becoming more reliable, but some still had downtime issues,” <http://www.networkworld.com/article/2866950/cloud-computing/which-cloud-providers-had-the-best-uptime-last-year.html>.
- [10] CloudHarmony Cloud status,<https://cloudharmony.com/status-1-year-of-storage-and-compute-group-by-regions-and-provider>.
- [11] Google Computing Engine. <https://cloud.google.com/products/compute-engine>
- [12] “HDFS architecture,” <http://hadoop.apache.org/docs/r2.2.0/hadoop-project-dist/hadoop-hdfs/HdfsDesign.html>.
- [13] Infosecurity-magazine, “Ddos-ers launch attacks from amazon ec2,” <http://goo.gl/vrXrHE>, July 2014, [Accessed September 25, 2014]
- [14] Microsoft Windows Azure. <http://www.windowsazure.com>
- [15] Rackspace. <http://www.rackspace.com/>
- [16] B. Russell: Passive Benchmarking with docker LXC, KVM & OpenStack, <http://www.slideshare.net/BodenRussell/kvm-anddocker-lxc-benchmarking-with-openstack>, retrieved February 2015.
- [17] D. Merkel, “Docker: Lightweight Linux Containers for Consistent Development and Deployment,” Linux Journal, vol. 2014, no. 239, p. 2, 2014.
- [18] J. Dean and S. Ghemawat, “Mapreduce: Simplified data processing on large clusters,” Commun. ACM, vol. 51, no. 1, pp. 107–113, Jan. 2008.

- [19] J. L. Berral, I. Goiri, R. Nou, F. Juli, J. Guitart, R. Gavalda, and J. Torres, “Towards energy-aware scheduling in data centers using machine learning,” in Proc. of the 1st Intl Conf. on energy-Efficient Computing and Networking, 2010, pp. 215–224.
- [20] K. Pepple, Deploying OpenStack. ” O’Reilly Media, Inc.”, 2011.
- [21] M. Mohaan and R. Raithatha, Learning Ansible. Packt Publishing Ltd,2014.
- [22] M. Taylor and S. Vargo, Learning Chef: A Guide to Configuration Management and Automation. O’Reilly Media, 2014.
- [23] T. Uphill, Mastering Puppet. Packt Publishing Ltd, 2014.