

行政院所屬各機關因公出國人員出國報告書
(出國類別：實習)

赴美國實習

『線路設備管理運用資訊二代系統(IOIS/eLEAMIS)軟體實習』

出國報告書

服務機關：中華電信股份有限公司
姓名職稱：李谷滿 助理工程師
服務單位：中華電信北區分公司
出國地區：美國
出國期間：自 92 年 9 月 21 日至 92 年 10 月 4 日
報告日期：92 年 12 月 31 日

系統識別號:C09203311

公務出國報告提要

頁數: 38 含附件: 否

報告名稱:

線路設備管理運用資訊二代系統(IOIS/eLEAMIS)軟體實習

主辦機關:

中華電信台灣北區電信分公司

聯絡人/電話:

盧婉屏/2344-3261

出國人員:

李谷滿 中華電信台灣北區電信分公司 客戶網路處 助理工程師

出國類別: 實習

出國地區: 美國

出國期間: 民國 92 年 09 月 21 日 - 民國 92 年 10 月 04 日

報告日期: 民國 92 年 12 月 31 日

分類號/目: G10/電子工程 /

關鍵詞: 公用數據中心,高可用度系統,UDC,HA

內容摘要: 本報告書主要分為目的、公用數據中心、高可用度系統及實習心得與建議、參考資料等五部分，兩項實習主題部分，首先針對公用數據中心 (Utility Data Center) 架構及組成單元予以說明，並介紹數據中心基本運作方式及特殊專用語說明，第二單元則介紹高可用度系統的原理與目前各主要高可用度系統比較分析，最後，就二項實習項目提出所獲心得及建議。

本文電子檔已上傳至出國報告資訊網

目 錄

一、目的	-----	2
二、公用數據中心實習報告	-----	2
三、高可用性系統實習報告	-----	23
四、實習心得與建議	-----	35
五、參考資料	-----	38

一、 目的：

IOIS/eLEAMIS 運用多層次分散式架構、為強化機房系統主機資源管理(Utility Data Center)、提升主系統與軟體程式最高可用度(High Availability)，以利協助線路設備管理運用資訊二代系統順利上線及提高新系統資源維運有效管理。

二、 過程：

1. 研習內容：

公用數據中心 (Utility Data Center) 實習報告

對於 IT 設備而言，為實現 IT 資源供應與業務需求變化保持一致，大大地降低 IT 成本，提高資源利用率和業務靈活性需要具備動態資源優化的能力，以實現虛擬服務器資源的協調，提供最廣泛的虛擬化解決方案最典型範例之一就是惠普公用數據中心

(Utility Data Center) 解決方案。該方案為可擴展、可編程的數據中心提供了體系結構和軟件，甚至還能夠使大型計算環境在保持始終如一優質服務的同時提高靈活性和經濟性。目前，該方案已經被一些企業所應用，通過 UDC，這些企業把 IT 資產虛擬為一個單一的邏輯實體進行管理，從而降低了 IT 複雜性並推動企業邁向公用化計算之路。

1.1 UDC 是一個支持業務靈活性的全面 IT 解決方案：

- 時間—HP UDC 內置的適應能力能夠即時回應服務要求，最大限度地減少人工干預。長遠來看，UDC 體系結構可擴展至數千台伺服器，實現幾乎無限的縱向擴充。
- 範圍—UDC 使您可以隨時隨地實現資源的物理和邏輯連接，跨越惠普和第三方元件以及不同的多作業系統傳統環境。
- 難易度—集成服務管理 (ISM) 將運營和業務支援功能與公用控制器軟體緊密集成，以實現無縫、經濟高效的資源管理。
- UDC 虛擬體系結構釋放並提供未充分利用的容量，可在需求分配出現極端波動的情況下預見並避免出現瓶頸。這樣，用戶可根據業務要求，快速無縫地重新部署高可用性、大容量計算資源。

1.2 UDC (Utility Data Center)公用數據中心對企業優勢：

1. 大幅降低操作成本百分之五十。
2. 在數據中心資源運用上卓越的進步。
3. 在資源部署與調度上靈活快速。

1.3 UDC (Utility Data Center)公用數據中心基本架構：

1. 公用控制器軟體 (Utility Controller Software)---負責新需求之資源計畫、設計、管理功能，此模組提供多廠家硬體支援、多重作業系統支援及多種儲存媒體支援，可彈性使用您所需要的

作業平台。

2. 管理單元(management rack)---

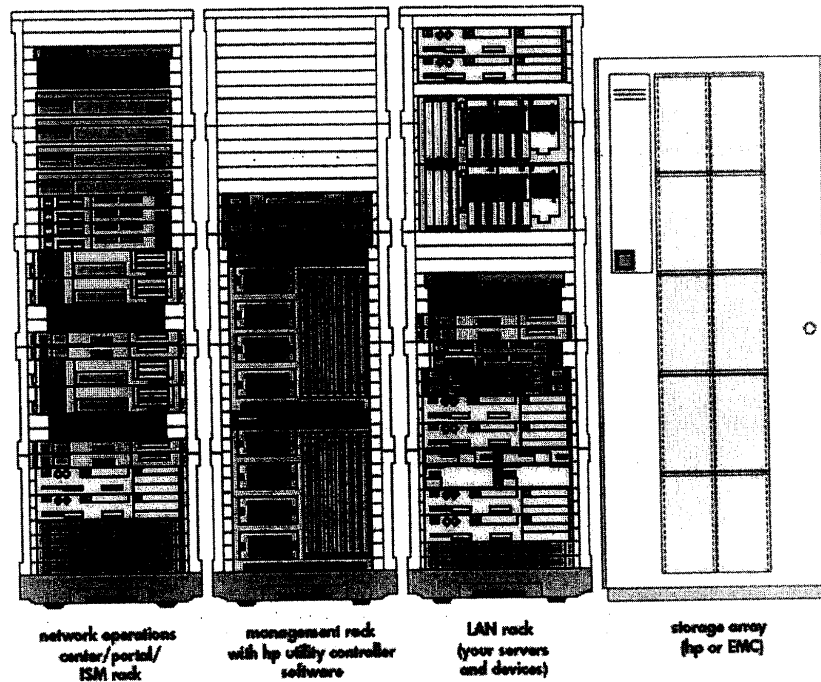
負責自動化了解數據中心拓普架構及資源使用情形。

3. 操作整合服務管理單元 (operations center/portal/Integrated Service Management rack)

4. 設備線路整合單元(linked racks of your servers and other equipment)

5. 資源計畫運用與優化軟體 (sophisticated capacity planning and optimization software)

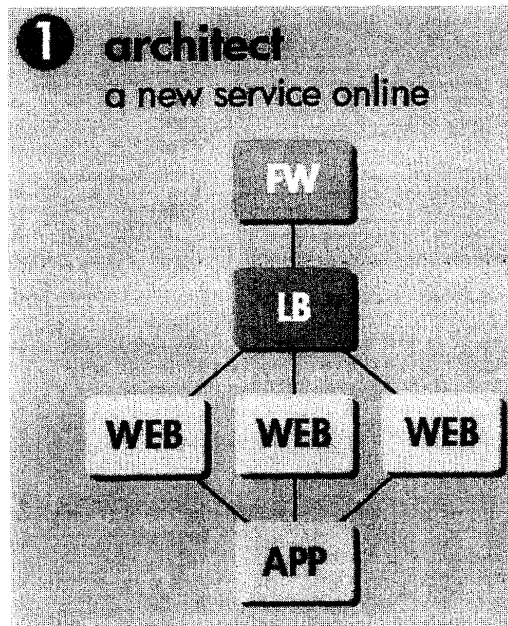
6. 儲存單元(storage arrays)



UDC 數據中心至少可管理 500 部以上不同廠商製造的主機伺服器與網路設備，並支援不同品牌之作業系統同時執行運作。

1.4 新服務或新系統加入 UDC 公用數據中心運作情形：

(1)當業務變更或新增營業項目時，此時業務需求與資源供應產生了變化，依據新業務功能要求，需新增網路與主機設備，防火牆(FW)壹部、負載平衡器(LB) 壹部、網路伺服器(APP) 壹部與應用伺服器(WEB) 三部才能達到新業務功能要求，設備架構如下圖---

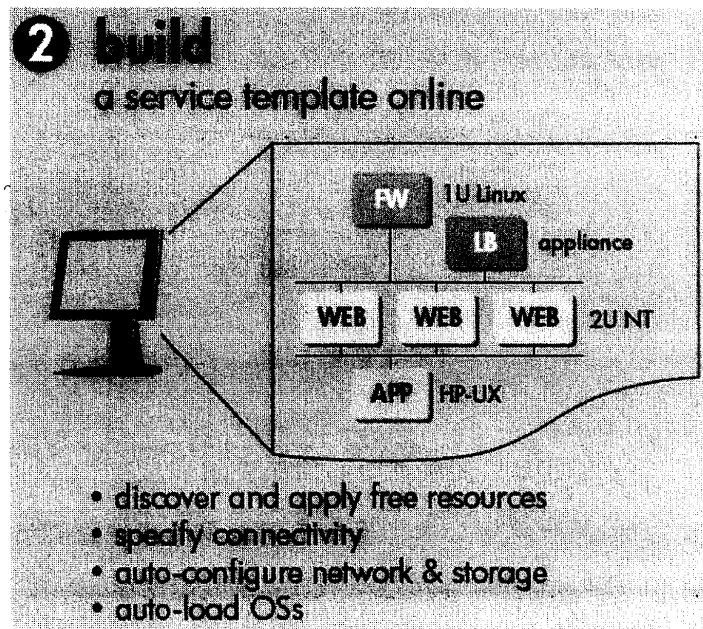


(2)當 UDC 數據中心接獲新增需求及了解新功能

所需設備時，執行下列步驟---

- 數據中心先將需求資源與目前可用資源運算並將所需求資源釋放出來。
- 設備間連線設定。
- 自動網路與儲存器設定。
- 依程式對需作業系統求，自動安裝作業系統。

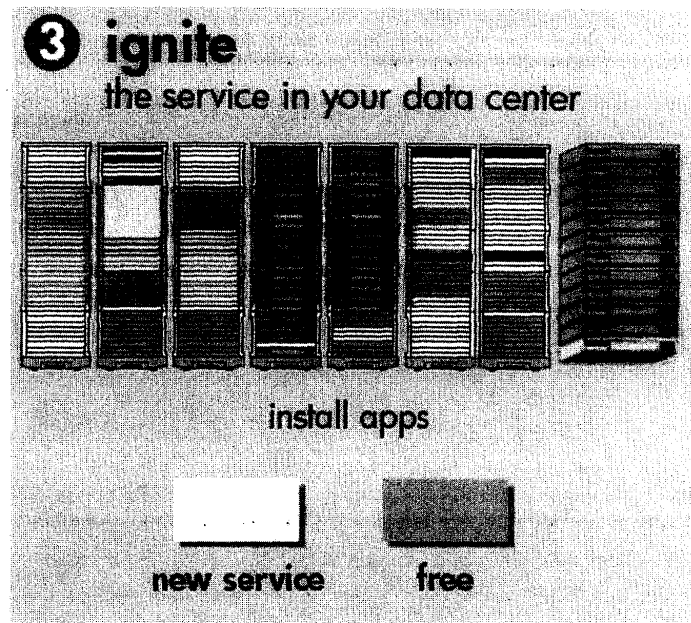
如下圖---



(3)在既有的設備上啓動並使用新服務項目，完全不用新購設備，大大地降低 IT 成本，快速靈活的新增功能，提高資源利用率和業務靈活性需要具備動態資源優化的能力，以實現虛擬服務器資源的協調。

UDC 數據中心創新架構允許

- 即刻完成基礎建設連線。
- 資源重新分配就如 drag-and-drop 般地容易

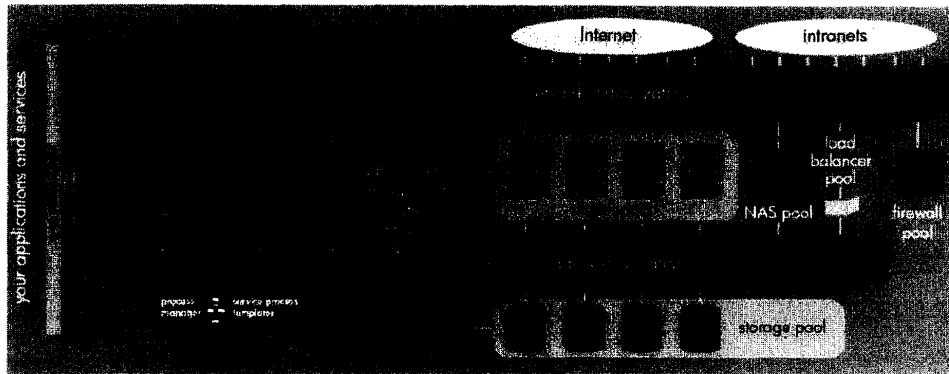


(完成新服務加入到 UDC 數據中心)

1.5 UDC 公用數據中心單元介紹：

公用數據中心最主要包含四個單元資料中心管理 (Data center management)、虛擬網路(network virtualization)、儲存媒體(storage virtualization)、客戶服務程式與系統介面。

(公用數據中心單元模組)



1.5.1 UDC 數據中心最主要特色之一是資源虛擬化

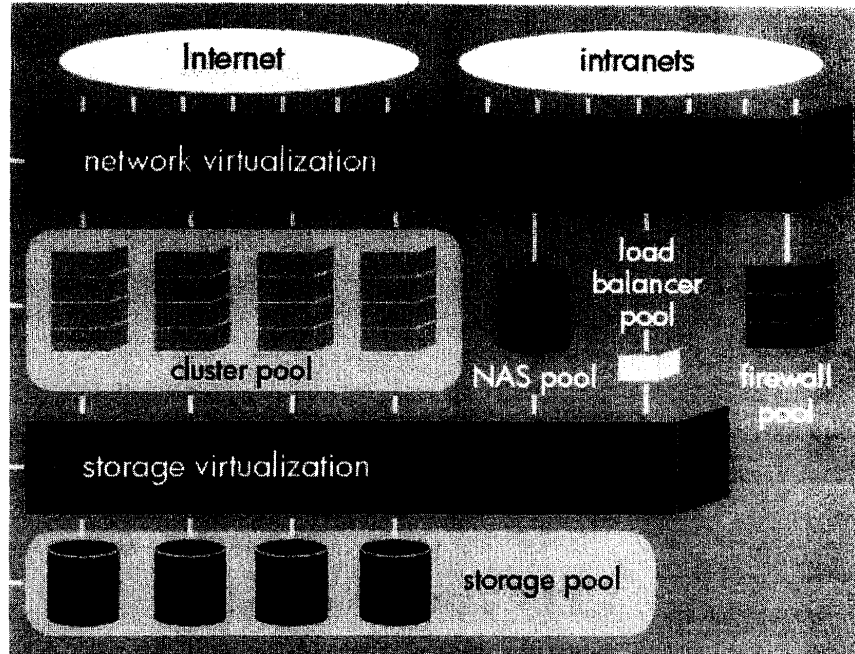
(resource virtualization) 。儲存媒體虛擬化包含儲存媒體與主機伺服器間界接網路，網路與主機則使用虛擬區域網路(virtual LAN) 。

在 UDC 數據中心兩種主要資源被虛擬化---

(1) 網路資源(network resources) ：

開創虛擬區域網路(virtual LAN) ，藉由程式化控制將主機與各項設備做虛擬重新接線，經由可程式化網路交換器連接到主機或從主機中移除，而達到虛擬接線(Virtual wiring)功能。

如下圖 network virtualization 部分



(2) 儲存媒體 (storage resources) :

藉由整份磁碟映像 (disk images) 而保有所有應用程式環境 (application environments)、檔案系統 (file systems)、可開機的作業系統映像 (bootable operating system images)、應用軟體 (application software) 等，主機可以透過 SCSI 介面從儲存體內找到機器開機的映像檔執行開機或取得其他資料。

如上圖 storage virtualization 部分。

1.5.2 UDC 平台對於新服務有兩個主要優點：

- 自動服務調度：藉由兩組分開資料儲存系統達成可調度式服務即可程式化調度與主機連

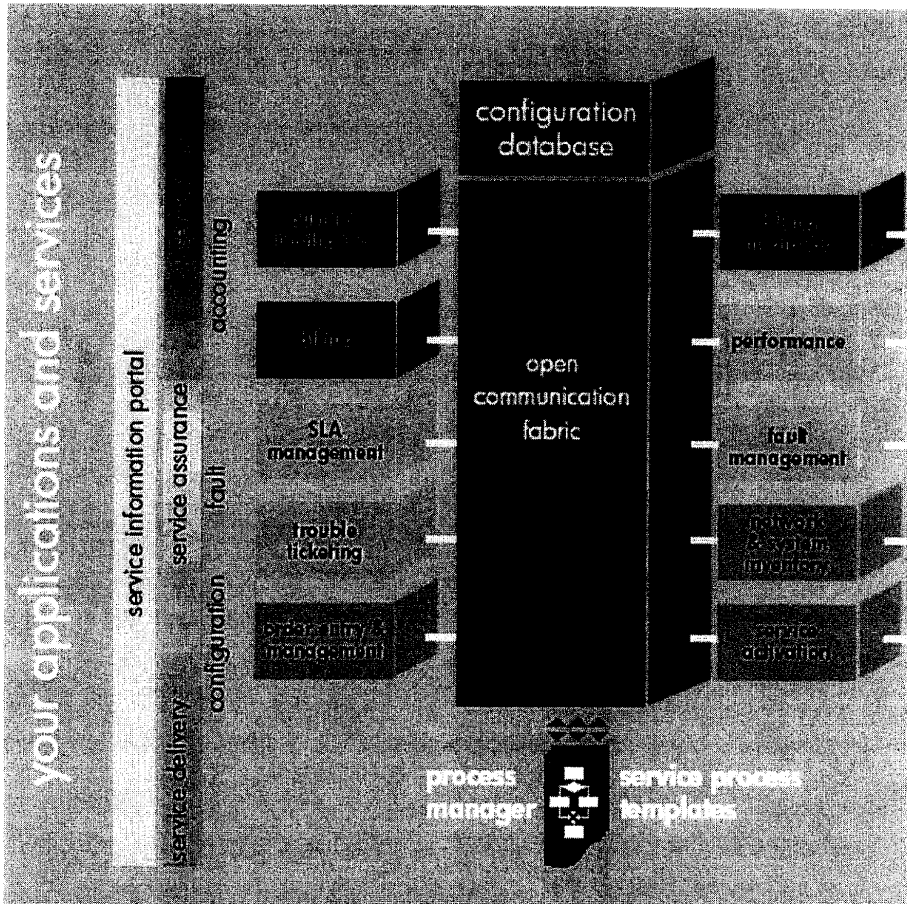
結。

- 動態服務容量能力調整：可動態調整服務內容及吸收外加需求，只要一經啓動設定即可使用。新服務單元是由空閒的資源池提供載入設定好的服務環境，透過虛擬連接線 (Virtually wiring) ，配上適當儲存體，再透過適當儲存體，安裝必須的作業系統，載入自行開發應用程式，新服務即可運用。

UDC 平台對於新服務有兩個主要優點：

- (1)自動服務調度。
- (2)動態服務容量能力調整。

(如下圖)

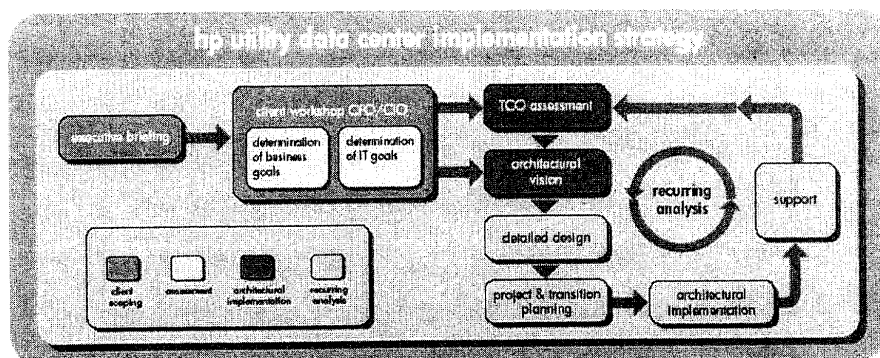


為支持動成長企業必須滿足以下三個要求：

- IT 資源必須以服務的形式交付：Web Services 正促使 IT 行業過渡到服務交付模式。在此模式下，基礎設施和 IT 服務，包括應用在內，將在需要的時間和地點交付、分配和支付。
- 資源必須是虛擬化的：虛擬化意味著通過對物理伺服器、存儲和網路的管理和控制、創建虛擬資源—計算、資訊和通信。虛擬化將顯著縮短適應時間，將其從購置或重新部署所花費的時間長度轉變為運行智慧管理軟體所需的毫秒量級。
- 端到端環境必須成為基於業務流程的環境：能夠智慧地回應和適應不斷變化的業務條件。這是適應性重要的核心新需求即：持續分析業務需求和智慧交付管理資源，以優化業務能力和靈活性。

介紹虛擬化服務器環境（VSE），VSE 能夠根據每個應用的服務目標和優先級來動態增長或減少每個應用所需要的資源配置，大大改善企業 IT 架構的業務靈活性和適應性，通過虛擬化，企業將可以創建動態的、自動化的資源，從而實現與業務需求變化保持一致的 IT 資源供應，最大限度地降低 IT 成本，提高資源利用率和業務靈活性。對於快速成長服務企業而言，需要具備動態資源優化的能力，以實現虛擬服務器資源的智能協調。VSE 的核心是智能政策引擎，通過智能政策引擎 VSE 可以管理資源池，從而實現虛

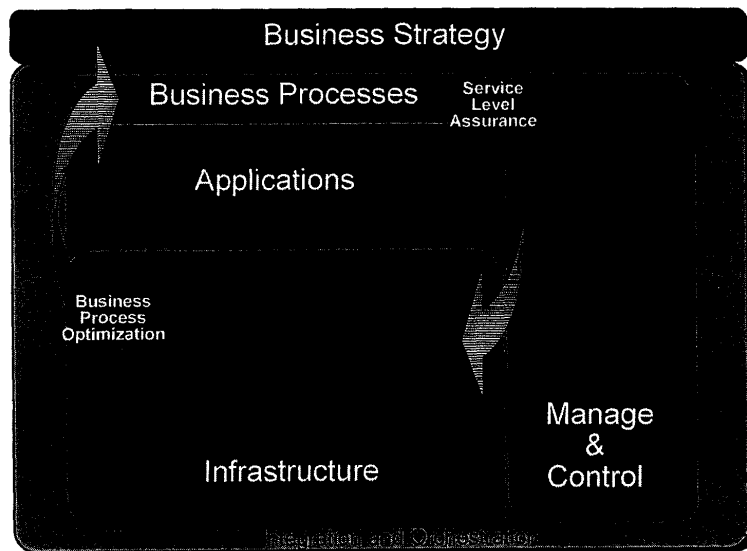
擬服務器資源的動態移動。HP-UX 工作負載管理軟件 (WLM) 是當前 Unix 領域獨有的基於目標的政策引擎，支持廣泛的服務器虛擬化以及靈活的垂直和水平擴展。當前 HP-UX 工作負載管理軟件只支持單一惠普服務器資源和工作流管理能力，用戶要實現服務器虛擬環境，通過工作復合管理軟件，再進行一些調整就可以投入使用了。未來，工作負載管理軟件還將實現跨平台的服務器群虛擬化管理，當發現一個服務器的能力不夠用了，通過網絡和軟件調用另一個服務器。HP-UX 工作負載管理軟件的價值在於自動性和職能兩個方面，過去是人來監控，當發現速度慢的時候，要打電話給他，再靠別的人來做，而現在是軟件監控然後自動做出調整，具有更強的可靠性、靈活性而且速度也更快。



典型範例之一就是惠普公用數據中心 (UDC) 解決方案。該方案為可擴展、可編程的數據中心提供了體系結構和軟件，甚至還能夠使大型計算環境在保持始終如一優質服務的同時提高靈活性和經濟性。目

前，該方案已經被一些企業所應用，通過 UDC，這些企業把 IT 資產虛擬為一個單一的邏輯實體進行管理，從而降低了 IT 複雜性並推動企業邁向公用化計算之路。他相信，隨著服務器虛擬技術的不斷發展和完善以及用戶對服務器虛擬技術的瞭解，將有更多希望減少 IT 開支的企業用戶樂於購買虛擬服務器解決方案，服務器虛擬產品市場將呈現旺盛的市場活力。

快速且難以預測的變化是當今業務環境的最大特點。靈活性業務起源於這種變化。一些企業故意引起市場變化，而另一些企業則在競爭對手之前對外部變化做出回應。這兩種類型的企業都依賴於快捷的回應速度來確保競爭優勢。動成長企業應支援快速的行動和反應。然而，更換或重新部署資訊技術（IT）資產所需的時間卻常常成為實現業務靈活性的障礙。



計算資源作為基於靈活和虛擬化基礎設施的服務提供給。當以這種方式管理和控制 IT 服務和資源時，即可迅速完成擴充和重新部署，並擺脫它們所採用的基礎設施的局限。這有助於打破限制業務回應能力的“IT 惰性”。

現在，靈活性業務要求 IT 能迅速持續地適應不斷變化要求，與此同時保持服務質量並控制成本和風險。惠普動成長企業戰略可降低 IT 惰性的影響，並採取立即和智慧行動來適應新的業務要求。動成長企業願景闡述了全新的理念：端到端管理，將來自合作夥伴和供應商的應用跨越企業邊界擴展和鏈結到用戶，構建和集成基礎設施。這一切都是以評估為第一步，以測量為後續的過程。惠普的方案支持業務靈活性而不必大規模替換現有的 IT 部件。實現動成長企業建立在惠普的一系列傳統優勢上：開放的系統和合作夥伴關係、異構基礎設施部件的管理和個性化的技術藍圖。

客戶可確保從他們目前所擁有的資產中獲得適應性優勢，實現最大限度的再利用並降低成本，同時實現新經濟週期、合併、激增或市場行動計畫的靈活性。

1.6 單位值：服務需求與容量的單位化過程 -----

- 1.6.1 一個重要需預先供給公用數據中心作自動化決定的資料，就是如何將新服務需求與容量能夠以標準方式表達，並且標準化演算法依決定作運用。
- 1.6.2 虛擬伺服器(“virtual server”)來表述，藉由伺服器分享(“server shares”)來表現出服務的級數與容量。
- 1.6.3 伺服器分享(“server shares”)代表一部伺服器設定所能接受最大服務容量，當服務要被安裝在伺服器時所需衡量的單位。
- 1.6.4 三組虛擬伺服器環境，包含伺服器內皆使用伺服器分享(“server shares”)來表示已配置資源與仍可用資源。
- 1.6.5 規格化伺服器內容量大小的方法是以基準點”盒外”(“outside-the-box”)的觀點來表示。
- 1.6.6 以基準點”盒外”(“outside-the-box”)的觀點來表示，取代以傳統內部實體參數如 CPU 個數，快取體大小，磁碟機與記憶體容量大小來表示，以基準點”外盒”(“outside-the-box”)來表示更貼近虛擬主機觀點。

1.6.7 以伺服器分享(“server shares”)方法來總合伺服器內部多樣化行爲模式，整合成爲一個簡要單位，並以(“server shares”)爲單位之數據方式呈現。

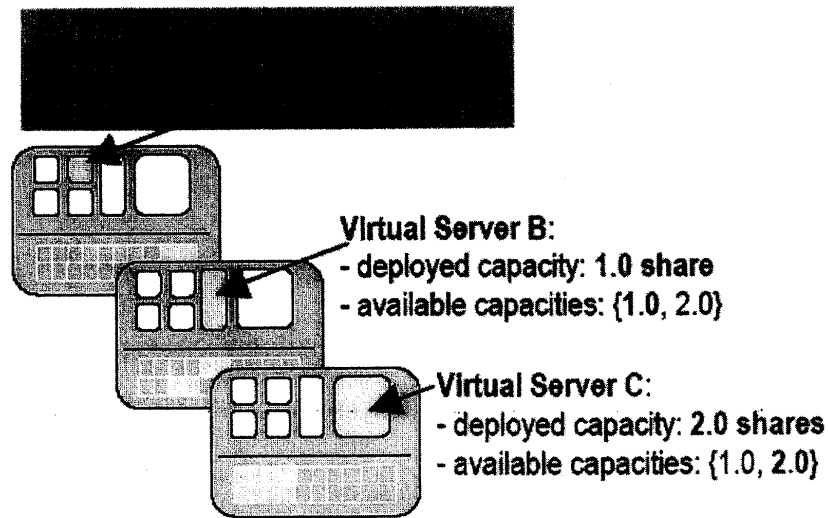
1.6.8 新服務需求(demand service)會被表達成爲相同服務的伺服器使用率，這就稱爲伺服器分享(“server shares”)。

1.6.9 已配置容量(*Deployed capacity*)：是指以指定資料中心的資源量。

1.6.10 下圖爲三個資料中心提供不同伺服器環境---

1.6.10.1 Virtual Server A：擁有一個已配置0.5 server share，而這0.5 server share 是從可用的伺服器容量{0.5, 2.0 or 10.0}環境中配置，目前有兩個資源{2.0 or 10.0}並未被使用，但之後可能會被Data center以控制指令啓動，這些未被使用資源會被統一對應到可用容量(*available capacities.*)。

1.6.10.2 Virtual Server B：擁有一個已配置1.0 server share，而這1.0 server share 是從可用的伺服器容量{1.0, 2.0}環境中配置，目前有一個資源{2.0}並未被使用，但之後可能會被Data center以控制指令啓動，這些未被使用資源會被統一對應到可用容量(*available capacities.*)。



1.6.10.3 Virtual Server C：擁有一個已配置2.0 server shares，而這2.0 server shares 是從可用的伺服器容量{1.0, 2.0}環境中配置，目前有一個資源{1.0}並未被使用，但之後可能會被Data center以控制指令啓動，這些未被使用資源會被統一對應到可用容量(*available capacities.*)。

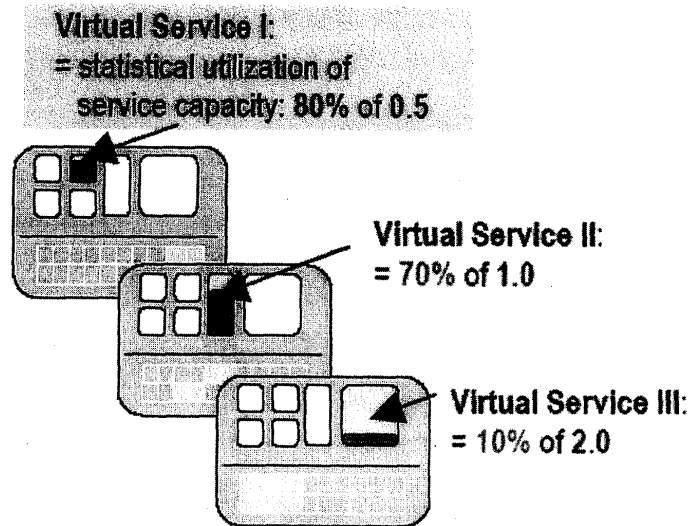
1.6.11 三個已配置資源的服務(deployed services)

1.6.11.1 服務需求是藉由服務分享(service shares)呈現出伺服器資源的使用情形。

1.6.11.2 虛擬服務一(Virtual service 1)：擁有 0.5 伺服器分享 (server share)的 0.8(80%)使用率(如下圖)。

1.6.11.3 虛擬服務二(Virtual service 2)：擁有 1.0 伺服器分享 (server share)的 0.7(70%)使用率(如下圖)。

1.6.11.4 虛擬服務三(Virtual service 3): 擁有 2.0 伺服器分享 (server share)的 0.1(10 %) 使用率(如下圖)。



1.7 把計算作為服務來提供和虛擬化是計算戰略歷史悠久的特點。適應性基礎設施技術和解決方案的新穎和獨特之處在於其協調與管理資源、流程和服務的方式。以下便是採用的適應性設計原則：

- 簡易化：複雜性（具有不同要求的多個獨立體系結構）是適應性的大敵。然而大規模更換並非理想選擇。通過管理和控制簡化現有基礎設施可以降低擴展或更換現有資源的風險。增添新組成部件時，選擇具有內在物理和邏輯簡單性的部件將能夠增強功能和適應性。

- **標準化**：標準可擴大簡化整個企業所帶來的好處，並簡化部署和使用 IT 資產的環境。
 - 介面標準化能夠降低通信開支並加快適應速度。
 - IT 服務標準化要求提升業務流程從基礎設施發出要求信號的速度，並降低了出錯率。
 - 標準化流程和策略，包括管理在內，加快資源供應情況的溝通以及資源本身的交付速度。

- **集成化**：消除系統部件之間的人為障礙可從利用率較低的資源釋放空間容量，確保整個 IT 基礎設施的互操作性。

- **模組化**：模組化適用於物理網路（伺服器和存儲設備）及其支援的邏輯虛擬資源。將 IT 資源作為獨立的邏輯設備有助於改善可管理性和回應能力。無需重新設計，便可增添新的容量。服務和合作夥伴也可被看作是擴展的基礎設施的模組化組成部件，用於在規模或趨勢變化期間管理最高負載。

1.8 利用適應性設計原則來為企業的基礎設施創建最靈活、高效和可靠的組成部件。這些部件包括伺服器、存儲設備、軟體和網路。同時，還可在每個部件中構建以下的核心能力：

- 連續和安全營運—能夠預測、診斷並快速響應潛在的和實際的故障條件。這為從單個伺服器子系統到整個企業提供一個了高度穩定的基礎設施。
- 自動和智慧管理—監視、維護、調節和控制硬體和資源。自動化智慧管理降低了資料中心甚至整個異構基礎設施的複雜性。最新管理工具使 IT 人員可以專心致力於戰略性業務規劃，而不必為日常管理工作而分心。
- 動態的資源優化—隨時隨地即時提供需要的計算資源。提供了出色的靈活性，充分利用當今及未來的資源、投資和系統。

在適應性基礎設施上實施的解決方案模式大大降低了 IT 複雜性，這是由於相對較少的獨特模式被映射到了相對較少的計算服務中。但在業務增長、合併或轉型時，可以獲得大多數好處。在簡化了業務流

程和所使用的 IT 後，新的業務環境可能只需改變規模。在這種情況下，根本不需要任何改變，因為適應性基礎設施和適應性管理解決方案及解決方案模式均具有可擴展性。如果有必要進行質變，那麼簡化業務流程與適應性基礎設施間的介面可提高進行各種變化的速度、經濟效益和成功機會。

2. 高可用性系統(High Availability)

在資訊科技蓬勃發展的環境下，企業間的競爭也變得更加劇烈。而系統擁有一個永不打烊的運作模式，更是一個不容忽視的競爭優勢。當資訊系統發生無可預期的停機事件時，企業損失生產力、貨料流失、客戶與收入。爲了修復系統而增加成本，還可能遭遇到進一步受罰、訴訟、聲譽不良的因此各式各樣企業對於高可用性系統的需求量也逐漸的在增加中。大部份的企業都希望系統能在上班時間提供可靠與連續運轉的服務。甚至有些公司需要對客戶提供 24 X 7 全年無休的服務。高可用性的電腦系統允許應用程式服務在硬體或軟體故障時繼續運作。希望透過本篇文章能更瞭解高可用性系統(High Availability)的架構、原理與如何去規劃一個永不打烊的運作模式。

2.1 何謂高可用性系統

高可用性系統(High Availability)一般來說是個對於系統的需求，甚至於是個對於整個環境而言的運轉力。對於不同的企業，高可用性系統代表著不同東西。可能是提供使用者正常上班時間外的支援。也可能是預防系統在上班時間發生意外的停機。各行各業的商務環境與需求，往往反映出它們對高可用性系統的期望：

全球性的服務

大量資料的處理

連續不休的運轉

生產力的管理

對一間有上述需求的公司而言，資訊系統等於是個關鍵系統(Critical System)不僅能處理資訊，它們還能幫忙管理事業。儘管這些系統的設計是可以信賴的，然而系統停機仍會對公司造成嚴重的影響，包括生產力耗損，客戶流失以及收入銳減。針對資訊環境，高可用性系統(High Availability)是不能單靠增加硬體來達到。那必須要靠加強一個環境(Environment)，處理(Process)，硬體(Hardware)與軟體(Software)的組合來減低意外停機事件的發生。造成無預警情況下停機，共有四大因素：

I. 硬體因素

ii. 軟體因素

iii. 環境因素

iv. 使用者因素

其中又以硬體因素所佔的比率最高(40%)，但是軟體與使用者因素有一直增加的現象。所以高可用性

系統(High Availability)的目的地就是減低這些因素發生的機率。

2.2 叢集高連續運轉方式

高連續運轉技術有幾種不同的做法。將介紹叢集(Cluster)高連續運轉方式，也就是使用 HP MC/ServiceGuard 技術。高可用性的電腦系統可讓應用程式服務在硬體或軟體故障時繼續運作。高可用性的系統讓使用者避免於軟體失敗和系統處理單元(SPU)，磁碟(Disk)或區域網路(LAN)元件故障。假如其中一個元件故障，其餘元件會接手監控。MC/ServiceGuard 和其他高可靠度子系統會協調元件間的轉換。

叢集(Cluster)是由伺服器的網路群組(稱為節點[node]的主機系統)，擁有備援(Redundant)的軟體和硬體，所以當任何一個關鍵元件故障時(也可稱為單點失敗[Single point of failure])，服務不會明顯地中斷。所以在架構上，高可用性的電腦系統必須要包括：

叢集元件的備援(Redundant Cluster Components)

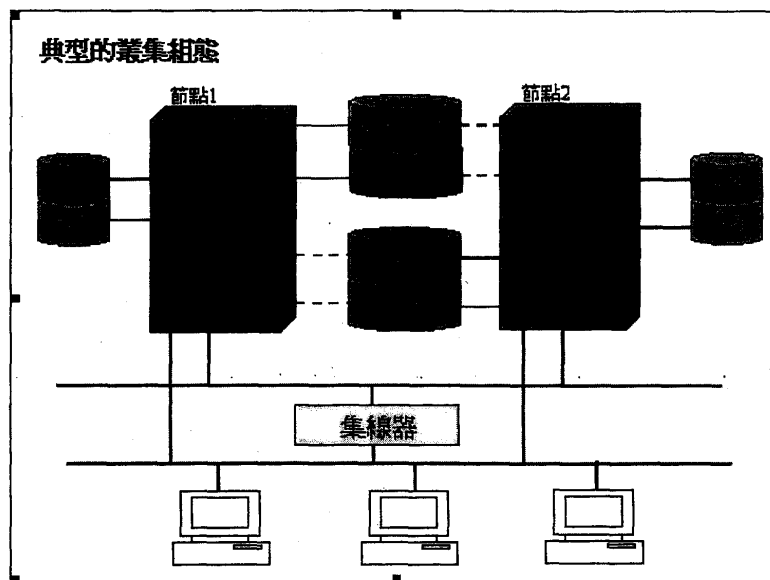
網路備援元件(Redundant Network Components)

備援磁碟儲存(Redundant Storage Space)

備援電源供應(Redundant Power Supply)

應用程式服務 (MC/ServiceGuard) 被群集在套件 (Package) 裡；當單一服務、節點 (Node)、網路或其他資源失敗時，能自動將套件的控制權轉移到叢集 (Cluster) 中的另一節點，盡量在不中斷的狀態下繼續服務。

如下圖展示目前使用 ha 的環境，有兩個節點的 MC/ServiceGuard 叢集。

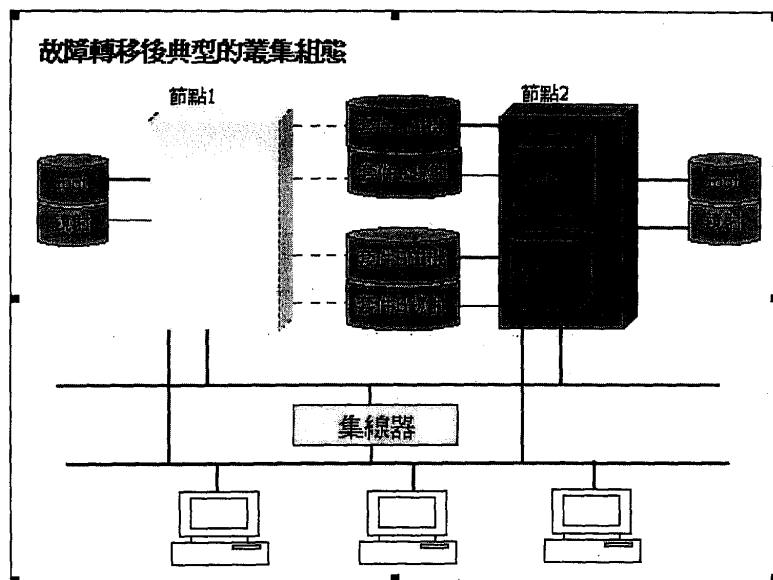


負責套件 A 的節點 1 和負責套件 B 的節點 2 都是各自獨立的系統處理單元 (SPU)。套件 A 和套件 B 都有各自需要的應用程式資料與鏡射資料磁碟群組。但是每個節點一定要和全部的鏡射資料磁碟群組串起來，而且一次只能有一個節點能存取套件資料磁碟群組。例如，套件 A 磁碟群組只能讓節點 1 用獨占的方式 (Exclusive) 存取，而套件 B 磁碟群組只能讓節點 2 用

獨占的方式 (Exclusive) 存取。但是節點 1 並不會對套件 B 磁碟群組做任何存取的動作。同樣的，節點 2 並不會對套件 A 磁碟群組做任何存取的動作。節點之間是靠冗餘的網路介面銜接起來，透過 TCP/IP 網路協定 (Protocol) 去傳遞訊息，包含心跳訊息 (Heartbeat messages)。

2.3 故障轉移 (Failover)

如圖展示的是故障轉移後的 MC/ServiceGuard 叢集 (Cluster)。



在正常情形下，MC/ServiceGuard 叢集 (Cluster) 會每隔一些時間透過心跳 (heartbeat) 訊息在套件各節點上執行時監控叢集 (Cluster) 元件是否正常。在 MC/ServiceGuard 叢集 (Cluster) 中執行的任何主機系

統稱為現行節點 (Active node)。建立套件時，您會指定主要節點 (Primary node) 與一或多個承接節點 (Adoptive nodes)。所以當網路連線或是節點故障時，MC/ServiceGuard 可以將套件的 control 權轉讓給下一個可用的承接節點。經過轉換之後，套件 (Package) 會繼續在承接節點上執行，一直到承接節點不能運轉為止。但是，您還可以選擇將套件配置成當主要節點再次上線後，讓套件回到主要節點上執行。

2.4 市場上有五種基本的資源群組失效接管策略，這些策略名稱因供應廠商的不同而異。

2.4.1 閒置備援 (Idle Standby)

主要節點擁有資源群組；備援節點閒置著，僅監督著主要節點。萬一主要主機故障，則由備援主機接管。這些節點是有優先次序的，意指有最高優先權的存活節點將取得資源群組。有較高優先權的節點加入叢集時將造成短暫的服務中斷。這些節點可以是不一樣強大的，備援節點只要夠強大到足以確保降級的運作。

2.4.2 輪替備援 (Rotating Standby)

兩個節點如同閒置備援一樣，但是沒有優先權之分。先進入叢集的節點將擁有資源群組，第二個節點則以備援節點加入。如果因為較高優先權的節點

加入而導致短暫的服務中斷是不必要的，這樣的失效接管策略顯得特別令人感興趣。如果叢集中存在多於兩個節點，輪替資源群組會變得相當混淆，加上這些節點在能力上必須大約同等。

2.4.3 簡易退卻 (Simple Fallover)

主要節點擁有資源群組。備援節點執行非關鍵性的應用程式（例如 web 或 FTP 伺服器，開發或測試環境）並將接管關鍵性的資源群組，但反之不然。如果備援節點不夠強大到能夠同時執行企業關鍵性的資源群組和非關鍵性的應用程式，後者可能被停止。一旦重新整合失效的節點，備援節點將釋放企業關鍵性應用程式的資源，並且潛在地重新啓動非關鍵性的負載。重新整合的主要節點接著重新要求企業關鍵性的資源群組。

2.4.4 相互接管 (Mutual Takeover)

基本上這是雙向的閒置備援：兩個伺服器被設定為兩者都能接管另一節點的資源群組。兩者都必須有足夠強大的 CPU 以足夠的速度執行雙方的應用程式 -- 或者效能的損失必須列入考量直到失效的節點重新整合。這在三個或更多的節點設定中也運作良好。

2.4.5 同步存取 (Concurrent Access)

所有的節點執行相同的資源群組（同步的資源群組中可以沒有 IP 或 MAC 地址！）並同時存取外部儲存體。目前市場上僅有一個應用程式使用到這個特徵：Oracle Parallel Server。這需要一個非常有效率而可能作為未來延伸的選項的鎖定管理程式。這裡的失效接管策略基本上是無效（null）的：如果一個應用程式實例失效，失效節點的 IP 地址可能會也可能不會被接管，取決於客戶端的理解力和適應性。

應用程式本身通常不是隸屬於 HA 軟體代理程式的監督。Linux-HA 應該是完全通用的，允許每個能夠在背景執行的應用程式 HA 而無須操作員介入（例如透過 cron）。可選擇地，可能有叢集 API & 函式庫作為未來的延伸，讓應用程式容易的和 HA 軟體代理程式溝通。

2.5 主流 UNIX 集群技術(Cluster)的高可用性產品比較分析-----

2.5.1 主要產品名稱

1. 康柏公司的 Tru64 UNIX 和 TrueCluster 服務器軟件。
2. 通用數據公司（Data General）的 DG/UX 和 DG/UX Clusters。
3. 惠普公司的 HP/UX 和 MC/ServiceGuard。
4. IBM 公司的 AIX HACMP。

5. Sequent 公司的 Dynix/ptx 和 ptx/CLUSTERS 。

6. Sun 公司的 Solaris 和 Sun Clusters 。

2.5.2 爲了盡可能全面、綜合地反映不同集群產品在爲用戶提供高可用性計算環境方面所具有的能力和特性，DHBA 的專家們根據企業計算用戶在實際應用中對集群系統的需求把整個分析過程劃分成了 6 個單項：

- (1) 集群備份與恢復。
- (2) 集群故障恢復配置能力。
- (3) 集群並行數據庫存取。
- (4) 集群高可用性管理。
- (5) 集群的單系統映像。
- (6) 災難恢復。

2.5.4 集群備份和恢復

「集群備份和恢復」能力，著重從故障恢復模式、檢測方法和恢復選項以及備份和恢復選項等幾個方面著手，同時需要考慮系統服務備份的分配、選擇本地與故障恢復的靈活性、並行與串行恢復，以及一系列諸如故障和資源短期等集群系統可檢測的條件。HP 的 MC/ServiceGuard 在這一單項中獲得最高分，除了擁有其他系統的共性外，HP 還支持動態備份選擇功能。HP 通過本地和故障

恢復來支持從 WAN 適配器故障中進行恢復的功能。

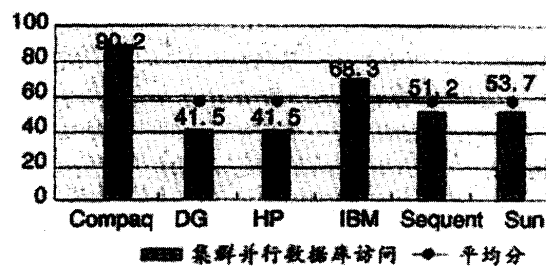
2.5.5 集群故障恢復配置能力

對「集群故障恢復配置能力」的考察，主要是針對系統故障恢復、自動配置硬件和系統軟件資源的能力。集群系統硬件配置能力決定於系統對於不同種類的服務器、磁盤和用於心跳和客戶機連接的網絡配置能力；集群軟件配置能力則主要決定於系統對混合操作系統版本（能夠在運行中升級）、可供選擇的中間件和 RAID 軟件的配置能力。本項目的考察重點涉及：節點/心跳媒體、RAID、中間件/NFS、TCP/IP 地址接管和磁盤互連。IBM 以 91.6 分獲得本項目的第一名。除了具備所有廠商共有的特性之外，IBM 在故障恢復配置中可支持 16 個以上的節點，從而獲得了很高的信任度。

2.5.6 集群並行數據庫訪問

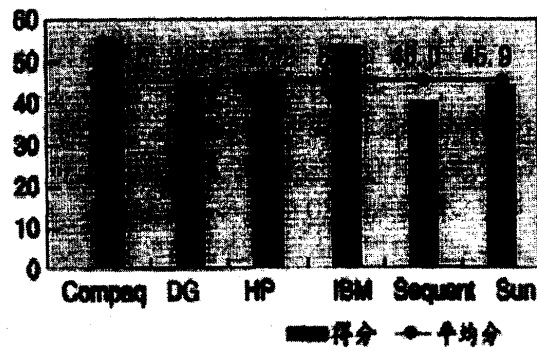
並行訪問允許多個集群節點共享一個物理數據庫。它要求 Distributed Lock Manager (DLM) 來協調並行的請求以更新共享的數據，這樣多個用戶就不會互相覆蓋彼此的更新。並行訪問可以加強可伸縮性，並可進行更快速的恢復，從而提升可用性優勢。康柏

的 TruClusterServer 軟件以 90.2 分的絕對優勢獲得了該項目的冠軍。TruCluster 在並行數據庫訪問配置中支持的節點多達 8 個。利用了虛擬原磁盤訪問，以及類似於康柏內存通道技術的高速節點互聯的 TruCluster，在提供並行數據訪問支持方面表現出了突出的優勢。



2.5.7 根據 DHBA 提供的研究數據，康柏公司的 Tru64 UNIX 和 Tru Cluster 服務器軟件以 53.5 分的最高成績，摘取綜合評分桂冠，並在「集群並行數據存取」、「集群高可用性管理」和「集群單系統映像」等三個單項評比中獲得第一名。IBM 以 51.8 分的總分位居第二。

在對 6 家廠商的 UNIX 操作系統和集群軟件的數百個特性和功能進行認真分析和評估後，所得評分(滿分六十分)。



集群備份與恢復加權值 20%、集群故障恢復配置能力加權值 20%、集群並行數據庫存取加權值 10%、集群高可用性管理加權值 5%、集群的單系統映像加權值 5%、災難恢復加權值 5%。

隨著各廠商在服務器、操作系統以及集群軟件等方面技術的不斷提升，各家廠商在以上項目測試中將會取得更加出色的成績。

三、實習心得與建議

1. 公用數據中心（Utility Data Center）方面：

- 1.1. 資訊系統實現業務靈活性要求，提供價值超出系統本身範圍以外的業務和規劃工具，並抓住快速變化的最新業務以更有效地部署公司內部現有資產。
- 1.2. 市場適應性（以服務為中心、虛擬、業務管理）提供適應性基礎主機相關設施計畫的早期實利，把我們的業務納入實現更大靈活性的軌道，並將系統構成任何擴展計畫的基礎核心。
- 1.3. 資訊系統結構有助於整合 IT 資產，更充分地利用現有基礎設施，節省更多管理時間和人員工作。
- 1.4. 資訊系統資源整合，可將既有系統設備整合為共用資源，成為一共用基礎建設，依資源主機效能需求，做適當配置與調整。
- 1.5. 面對市場快速變化需求，新業務、新服務不斷增加，可透過公用數據中心（Utility Data Center）的可編程、可擴展性，在為增加新設備下完成新服務需求。

2. 高可用性系統(High Availability)

- 2.1. 規劃和導入一個高連續運轉系統是會牽涉到整個 IT 環境的運作。如果沒有正確的完成設計上的每一個步驟，將不能達到預定的保護措施與效能成果。建議最好從規劃與導入原始的解決方案，到監控與必要的變更，於完善計畫後，並能確實達到原始需求，才進行採購。
- 2.2. 高可用性系統(High Availability)規劃有備援主機運作時，應於定時測試其故障轉移(Failover)功能是否能順利移轉，運作功能是否正常，並增加系統維護人員對高可用性系統(High Availability)維運經驗。
- 2.3. 本資訊系統採用 HP MC/Service Guard，對於故障轉移(Failover)期間，用戶連線將會中斷，直到故障轉移(Failover)完成，再重新連線，所以用戶將於故障轉移(Failover)期間無法使用，這也是目前高可用性系統(High Availability)待加強地方。由於此次參訪 HP Cupertino center，於他們研討會中介紹了 MC/Service Guard 此項產品，本人提出該公司產品的故障轉移(Failover)中斷連線問題，得知目前該公司針對此項功能已有解決策略，並已研發完成此軟體，此軟體稱為 HA TCP。
- 2.4. HA TCP 為一可保有用戶連線軟體，於主機故障轉移(Failover)期間用戶不中斷連線，目前並不是一個獨立軟體，HP 將他整合於 OpenCall SS7 產品內。

2.5. HA TCP 目前功能正在加強中，而 HP 原廠也於申請專利權後，期盼能有機會規劃為獨立產品，便可應用於一般 hpux mc/sg 上使用，若本公司內對於有新 hpux mc/sg 需求或已有該項產品單位可多注意。

參考資料：

- 1) UNIX 集群功能高可用性競爭分析報告(D.H.B.A)。
- 2) 高可用性系統之簡介(HP)。
- 3) IDC White paper(HP)。
- 4) Control Architecture for Service Grids in a Federation of Utility Data Center。
- 5) UDC SERVICE brief introduction。
- 6) Hp HSU, WAN-YEN (HP-Cupertino) information。