

第四章

資源管理與允入控制

4.1 預購法頻寬規劃

以預購法規劃頻寬之程序為：各 Ingress Router 中的 Bandwidth Order Agent(BOA)向 Bandwidth Broker (BB)批購資源供實際使用。批購的內容包含指定 link quality(以 quality entropy 表示)與其所需頻寬，BOA 據此啟動 bandwidth order procedure 向 BB 發送 resource request message 以取得資源，而 BB 則會依照各個 BOA 需求情況，考量如何進行資源分配。在選取 link 方面，BOA 會根據 Concerned Time Period 相關之 Referenced Time Period 找出該 Ingress Router 在 Reference Time Period 所選用過的路徑(由 LPPA 提供)，以這些路徑上所經過的 link 當作選取目標，以 link 為單位一一向 BB 批購資源。

4.1.1 Historical Traffic Pattern

預購所需之資源預測需要利用歷史資料，如果使用者需求具有重複性，即可用來預測。首先對每個不同時段，將 traffic 在各時段之 bandwidth 統計出來(圖 4.1)，根據使用者的行為不同，對於頻寬的需求也就不同，例如在離峰時段頻寬的需求量小，在尖峰時段頻寬的需求量大。頻寬需求在不同時段所顯現出來的差距較為明顯，以時段方式區分較具代表意義。以圖 4.1、4.2 為例，圖 4.2 為相同時段(亦為圖 4.1 對 x 軸作切割)，不同日期，在該時段之 bandwidth 紀錄。依照圖 4.2 之紀錄結果進行統計分析，分佈的狀況會依照時段不同而有所區別(圖 4.3)，如平均值(mean)、標準差(deviation)與範圍。

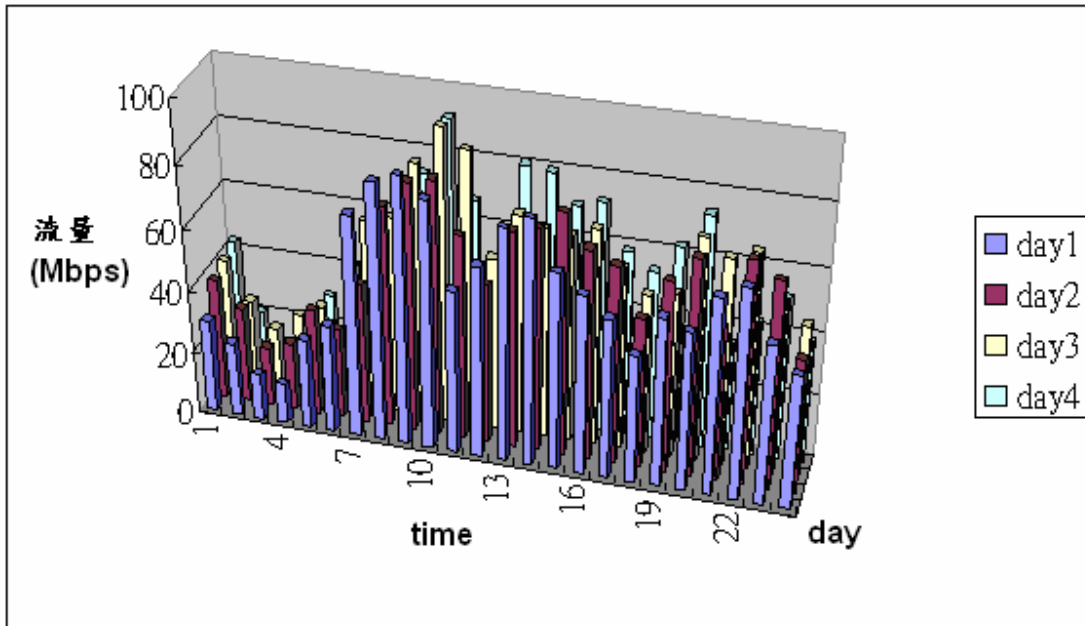


圖 4.1 : 以時段作為區分的流量統計

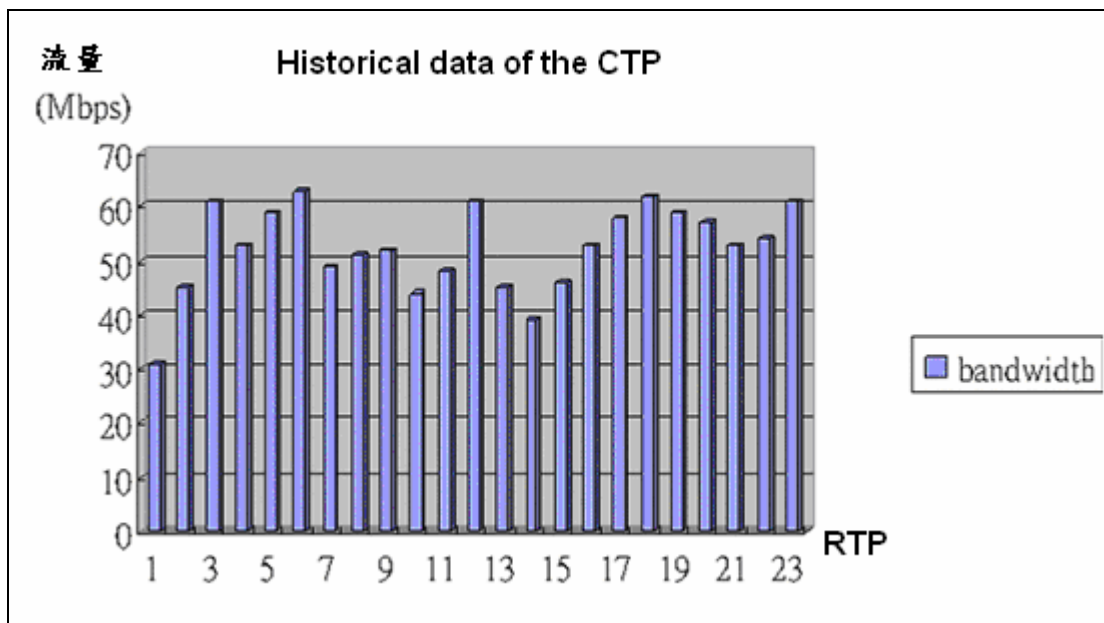


圖 4.2 : Bandwidth demand at RTPs with respect to a CTP

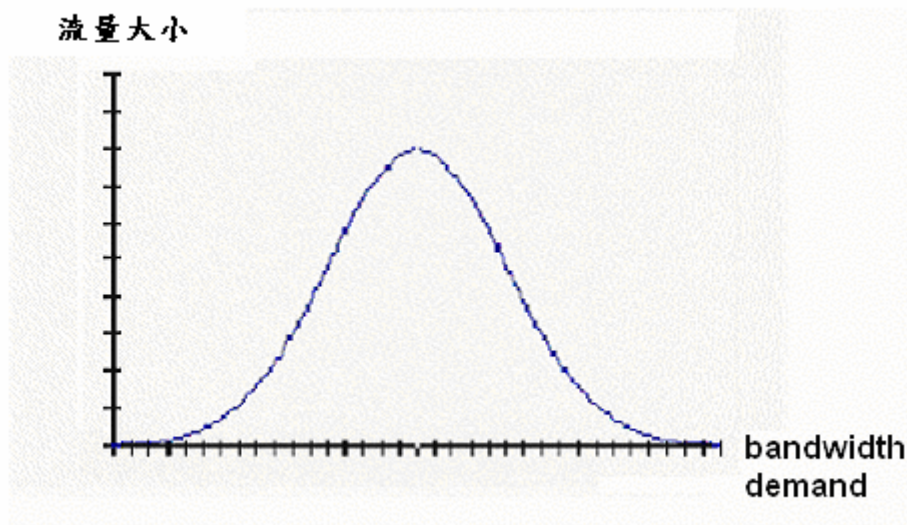


圖 4.3：Demand distribution at RTPs with respect to a CTP

4.1.2 最佳預購頻寬

根據過去不同時段之 traffic pattern、link 之 bandwidth distribution 與 Bandwidth Broker(BB)索費方式，BOA 以此針對各 link 預估出相對應的最佳頻寬批購量 θ (圖 4.4)。我們假設 BB 對於預先批購之單位頻寬價格為 C_1 ，而對於 Real-Time on-demand 批購之單位頻寬價格為 C_2 。

由於我們無法準確預測未來該 CTP 於執行時段之實際頻寬需求，若是採用保守的方式批購頻寬，雖然於預先批購頻寬階段以較低的成本預購頻寬，但是於執行時段易發生不能夠應付超過預期的頻寬需求而尋求較高成本的臨時批購 (on-demand)方式解決頻寬不足的問題，並且臨時批購所產生的訊息(message)交換，當數量到達一定程度時，會造成 BB 附近的網路須負擔額外的負荷。若是採用大膽的方式預先批購較多頻寬，首先於預先批購階段就得付出較多的頻寬成本，雖然在執行時段較不易出現臨時批購的現象，但是由於批購過多，當執行時段所需頻寬少於預先批購量時，即造成網路資源的浪費。(圖 4.5)。

因此 BOA 批購頻寬時當力求精準，避免預購太多(超過實際所需)造成頻寬

浪費，或批購不足以至於必須以較高價格臨時批購所需頻寬。

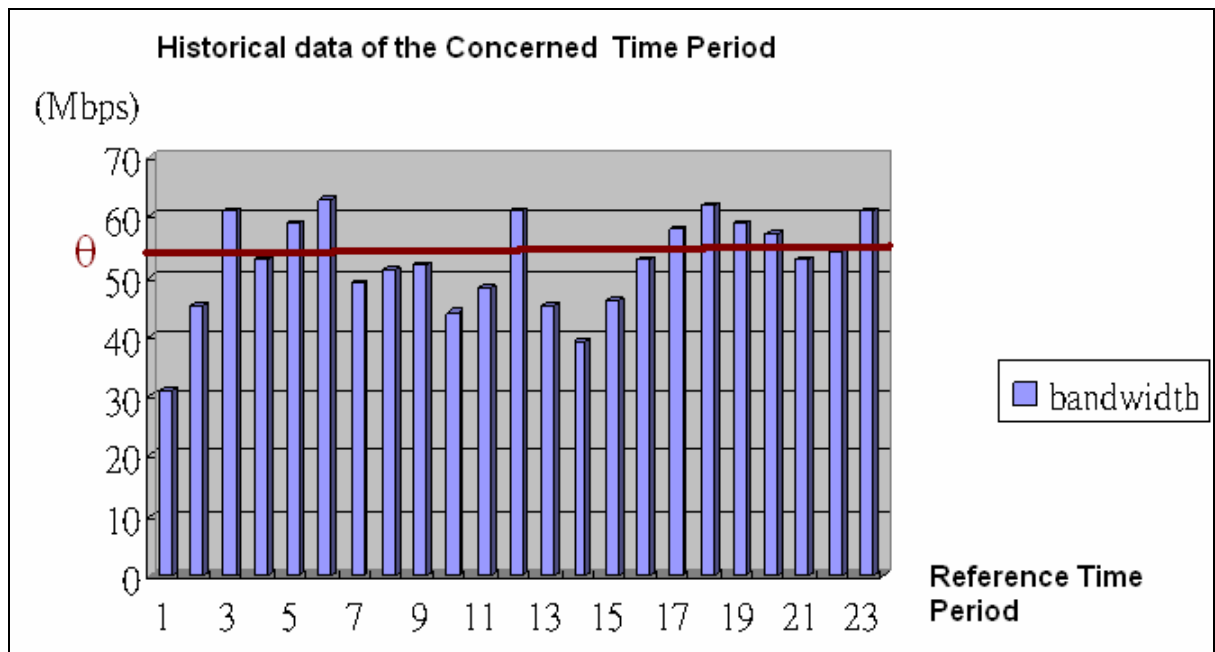


圖 4.4：根據歷史資料決定 θ ，用來當作預先批購的頻寬

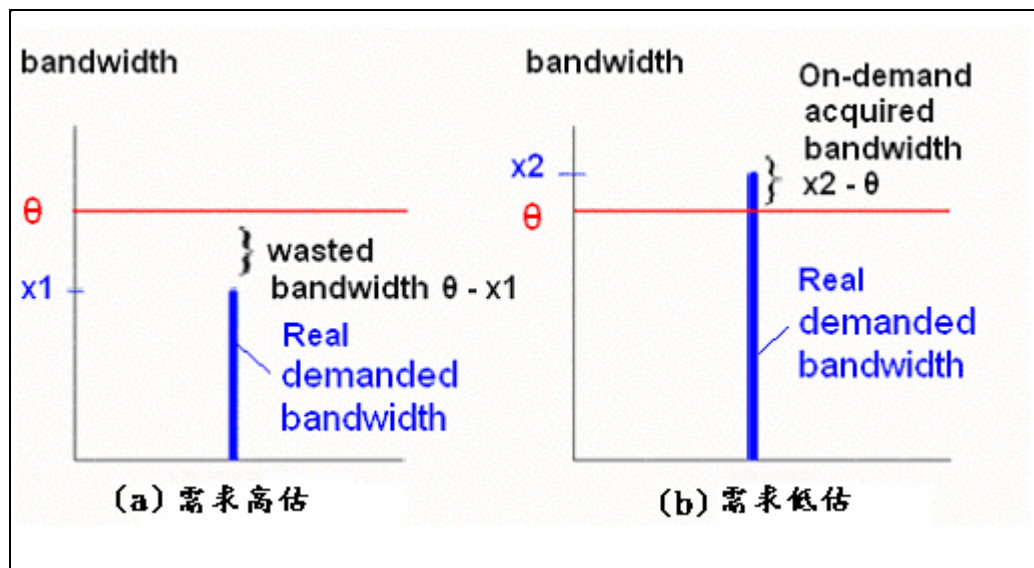


圖 4.5：預估需求與實際需求之差(a)需求高估
(b)需求低估

4.1.2.1.最佳化模型(Optimization Model)

由於用 RTP 的需求歷史預測未來的流量需求有其不準度，假設 CTP 的流量需求機率分佈與 RTP 流量統計分佈一致，本研究試圖找出成本期望值最低的最佳頻寬批購量的計算公式，納入機率分佈特性，根據過去不同時段的 traffic pattern、bandwidth distribution 與 Bandwidth Broker(BB)費率規則，可求得批購價格之最佳期望值並估計出每個時段適當的頻寬批購量 θ (如圖 4.6)。

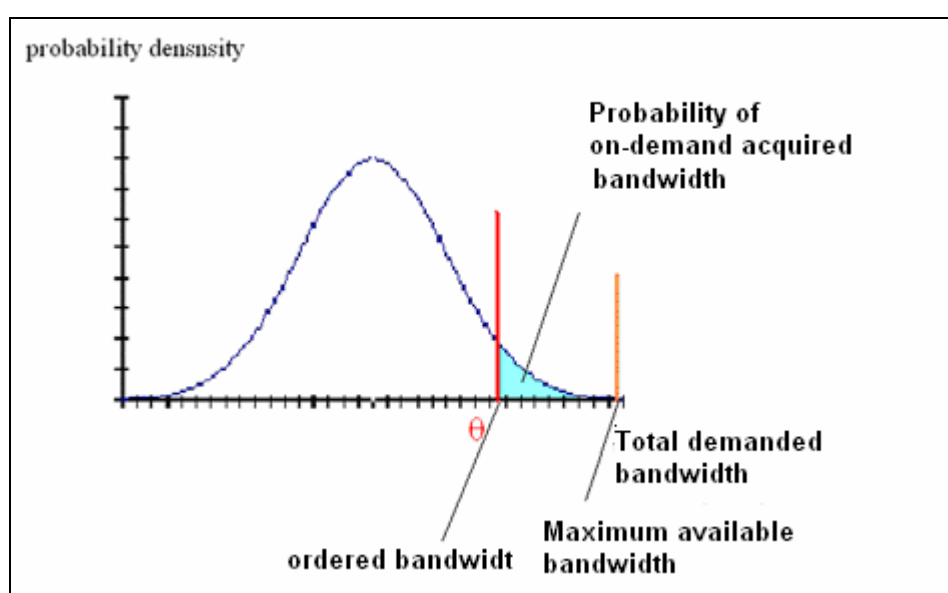


圖 4.6：最佳頻寬批購量 θ

如圖 4.6 所示，用過去各個 RTP 之頻寬曲線統計分佈作為 CTP 頻寬需求之機率密度，圖 4.6 以 normal distribution 作為例子。

我們以 $E\{C(\theta)\}$ 表示期望成本(expected cost)，期望成本分為兩個部分，第一個部分為預先批購成本，由預購頻寬量乘上 Bandwidth Broker(BB)對於預購所索取之成本 C_1 。第二個部分為 on-demand request 成本，由每次因為頻寬不足所補足的頻寬量乘上 BB 對於預購所索取之成本 C_2 ，在實際執行時段累計這些即時批購成本成為該時段所付出之 on-demand request 成本。將預先批購成本與 on-demand request 成本總合起來即為望成本(圖 4.7)。

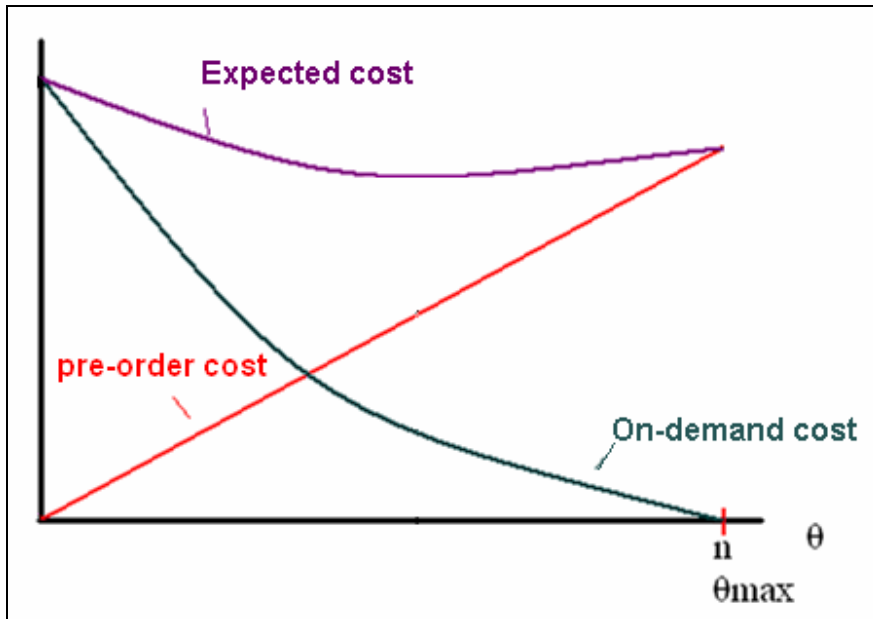


圖 4.7 : Expected cost、 pre-order cost and on-demand cost

Parameter List

- C_1 Pre-order request unit cost
- C_2 On-demand request unit cost
- i Bandwidth demand at CTP
- P_i Probability of demand bandwidth i
- θ Optimal pre-order bandwidth request
- $E\{C(\theta)\}$ Expected bandwidth cost

每個 CTP 之期望成本($E\{C(\theta)\}$)為 pre-order cost 與 on-demand cost 之總合，
如 Eq. 4-1(a)表示。

$$E\{C(\theta)\} = C_1 * \theta + C_2 * \int_0^n (i - \theta) * P_i di \quad (4-1 a)$$

由於 Eq. 4-1(a)必須在所使用的頻寬機率分佈 P_i 為連續時才適用，但是關於頻寬之統計方式有可能為了方便起見，採用不連續的方式(discrete)作紀錄，所以我們將 Eq. 4-1 稍作修改成 Eq. 4-1(b)，以供不連續之頻寬紀錄方式計算最佳批購值。

$$E\{C(\theta)\} = C_1 * \theta + C_2 * \sum_{i=0}^n (i - \theta) * P_i d_i \quad (4-1 b)$$

Eq. 4-1(a)、4-1(b) 中第一項是預購頻寬的成本，第二項是臨時批購的期望成本。

圖 4.8 為舉例表示($E\{C(\theta)\}$)的組成關係。

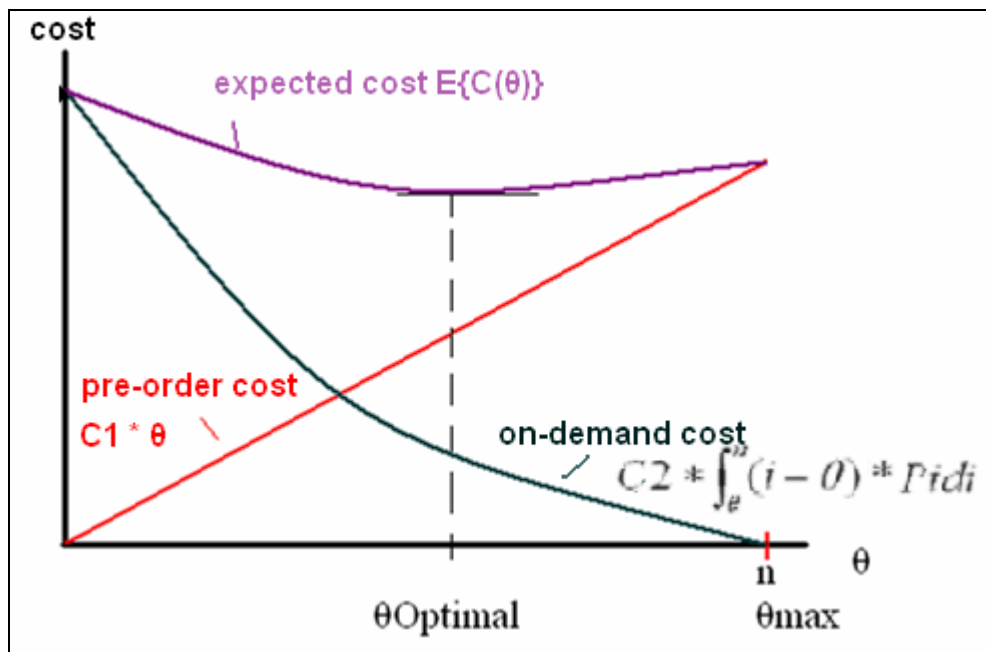


圖 4.8 : Expected cost、pre-order cost、on-demand cost and optimal θ

當頻寬 i 小於預先批購量 θ 時，於執行時段(run-time)預先批購的頻寬足以應付在該時段所使用的頻寬 i 而不需進行 on-demand request，在臨時批購成本皆為零，因此 Eq. 4-1(a)與 Eq. 4-1(b)中的臨時批購成本可化簡成 Eq. 4-2(a)與 Eq. 4-2(b)。

$$ondemand _ cost = C_2 * \int_{\theta}^n (i - \theta) * P_i d_i \quad (4-2 a)$$

$$\text{ondemand}_{\text{cost}} = C_2 * \sum_{i=\theta}^n (i - \theta) * P_i \quad (4-2 \text{ b})$$

Eq. 4-1(a)、Eq. 4-1(b)可簡化為 Eq. 4-3(a)與 Eq. 4-3(b)

$$E\{C(\theta)\} = C_1 * \theta + C_2 * \int_{\theta}^n (i - \theta) * P_i \quad (4-3 \text{ a})$$

$$E\{C(\theta)\} = C_1 * \theta + C_2 * \sum_{i=\theta}^n (i - \theta) * P_i \quad (4-3 \text{ b})$$

如果 Eq. 4-3(a)中的 $i * P_i$ 與 P_i 皆具有反導函數，最佳頻寬批購量可計算而得如下：

推導過程

$$E\{C(\theta)\} = C_1 * \theta + C_2 * \int_{\theta}^n (i - \theta) * P_i$$

$$E\{C(\theta)\} = C_1 * \theta + C_2 \left[\int_{\theta}^n i * P_i - \int_{\theta}^n \theta * P_i \right]$$

$$\frac{dE\{C(\theta)\}}{d\theta} = C_1 + C_2 \left[-\theta * P_{\theta} - \int_{\theta}^n P_i + \theta * P \right]$$

$$\frac{dE\{C(\theta)\}}{d\theta} = C_1 - C_2 \int_{\theta}^n P_i$$

$$\frac{d \frac{dE\{C(\theta)\}}{d\theta}}{d\theta} = C_2 * P_{\theta}$$

根據上述之推導過程，我們可以導出當 $C_2 < C_1$ 時最佳頻寬批購值之數學解 Eq. 4-4，並且推導可知預購法恆有最低預期成本，所以 BOA 可根據最低預期成本所對應之最佳預購頻寬值向 BB 預購。

$$\frac{C_1}{C_2} = \int_{\theta}^n P_i \quad (4-4)$$

4.1.2.2. 索取費率與預購頻寬

其次我們就 C_1 與 C_2 比例討論頻寬預購值 θ 之決策。由於不同的 Bandwidth Broker(BB)可能對於 pr-order cost(C_1)與 on-demand cost(C_2)有不同的制定原則而影響兩者之間的比例，若是 $C_2 < C_1$ 則臨時批購成本較預購成本低，BOA 於預購階段不會採取任何預購的動作，在執行時段對於每個進入訊務進行臨時批購取得頻寬允入；若 $C_2 > C_1$ 則臨時批購成本較預購成本高，假設 C_1 與 C_2 之比例相近，於預先批購階段批購較保守估計所需頻寬，執行時段超出預估的部分則採取臨時批購(on-demand request)會比較有利，如此可避免因為批購過多頻寬浪費成本；若是 C_1 與 C_2 比例相差懸殊，代表臨時批購成本偏高，則宜於預先批購階段批購較多，減少於實際執行階段發生臨時批購的機率，以降低執行時段之預期成本。隨著臨時批購成本越高，使得 Bandwidth Order Agent 在進行頻寬批購時，會預先批購較多的頻寬，也就是從 Eq. 4-4 中可得知 θ 會越靠近 n (maximum available bandwidth)。

當 Eq. 4-3(a)之 $i * P_i$ 與 P_i 其中任一變數不具有反導函數，不適用積分時，只能用 Eq. 4-3(b)，而 Eq. 4-3(b)則可用簡單的計算方式求出最佳頻寬批購值。

4.1.2.3. 多重服務等級之預購方案

由於我們以 Diffserv domain 作為核心網路，在 Diffserv 架構之下，某一 link 之使用可能分成數個等級，如 EF、AF1、AF2...、BE。當某一 link 頻寬有等級區分時，本研究所提出的 BOA 最佳頻寬批購之計算只需稍微修正即可沿用，我們可將每一個不同等級之頻寬於同一個 link 上視為不同各別單獨的 link，再套用所提出的最佳頻寬計算公式即可。

4.1.3 預購資源不足之解決方案

BOA 向 BB 批購頻寬，但是 BB 不一定有足夠頻寬滿足所有 Ingress Router 的批購。當 BOA 所獲得之分配頻寬低於預期時，BOA 可選擇不再批購替代資源，等到實際執行時段(Execution Time Period)將資源不足的部分以臨時取得頻寬的方式解決不足的問題，或者提前採取因應措。如果容許提升服務等級時，即可批購替代頻寬，將某個服務等級(service class)之路徑上不足頻寬之 link 根據需求頻寬差額，批購較高等級的頻寬替代，使得路徑上之各個 link 皆可滿足需求，以維持服務品質。

圖 4.9 為批購檢查流程，當 BOA 收到來自於 BB 所分配的頻寬(resource)時，先檢查是否符合預購量。若符合，則繼續完成批購程序，將資源交由 Local SPPA 依照各 link 上之可用頻寬進行路徑規劃。若 BB 所分配的頻寬少於預購的頻寬 θ ，則向 BB 要求次一等級頻寬作為替代。當 BB 各項服務等級的資源皆無法滿足 BOA 在預先批購階段以資源替代的方式對於預購不足之服務等級事先取得足夠的資源，則將問題延後到實際執行時段(Execution Time Period)，等到真正發生資源不足時，再採取臨時批購成本或是資源重新分配(resource reallocate)來補救。

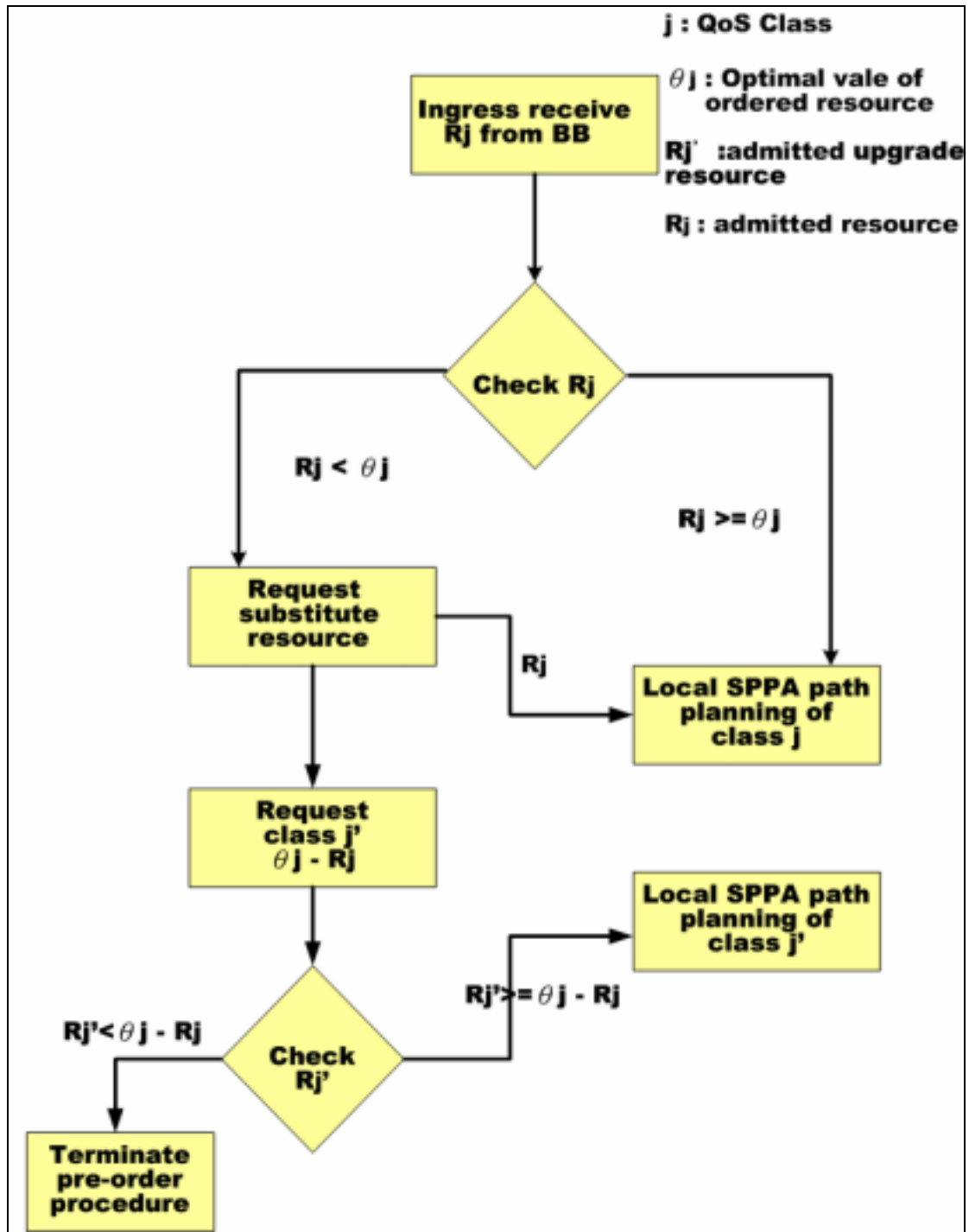


圖 4.9：頻寬預購流程

4.2 執行時段(Execution Time Period)之允入控制與資源管理

4.2.1 允入控制流程

經過事前的資源規劃，由 Bandwidth Order Agent(BOA)依照不同服務等級(service class)之預測結果向 Bandwidth Broker(BB)批購資源，其後將所批得的資源由 Path Planning Agent(PPA)規劃路徑，為不同服務等級之服務需求組合出適當的可選用路徑。Admission Control Agent(ACA)將這些具有不同服務品質的路徑選擇適當者分配給允入之服務需求。當某個訊務要進到某個 BBQ 核心網路時，ACA 會依照當時所能掌控的路徑資源作為允入的依據。在得知該訊務所需的品質需求後，若該服務等級的路徑所餘頻寬足夠，則允入該訊務，反之若該服務等級之資源不足，則尋求替代資源，當無替代資源時，則拒絕該訊務之進入，圖 4.10 為允入控制流程示意圖。

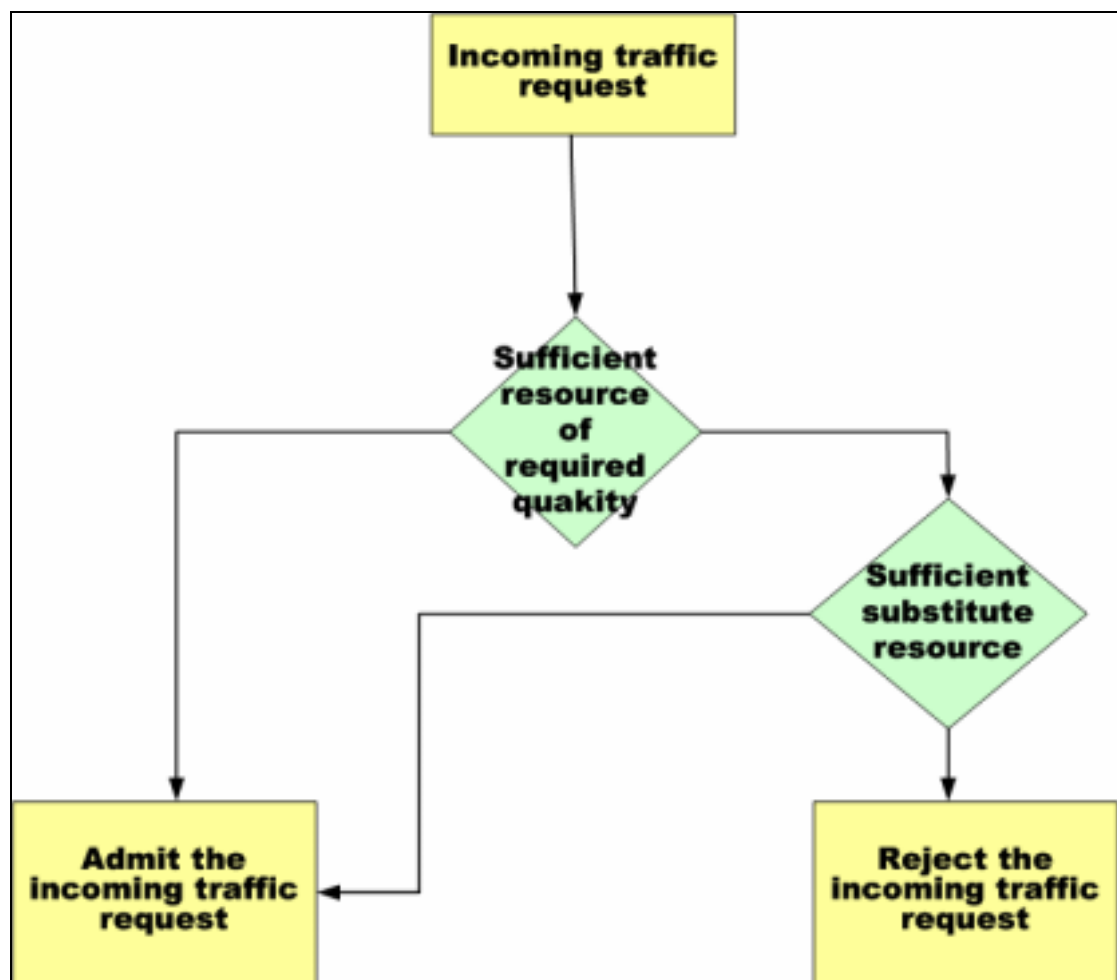


圖 4.10：允入控制流程

4.2.2 執行時段資源管理

執行時段(Execution Time Period)為之前各個預購資源規劃階段之 CTP 進入到實際執行階段，開始陸續出現提出資源需求之訊務，由 Ingress Router 負責接收需求訊息。Ingress Router 內部資源管理由 Bandwidth Order Agent(BOA)與 Admission Control Agent(ACA)共同負責，BOA 負責預購頻寬，ACA 則是負責允入控制。由於預購頻寬是根據 Bandwidth Broker(BB)之索費方式與過去統計的結果，選擇一個較適當頻寬預購值 θ ，經由 local SPPA 路徑規劃成可用之路徑資源交給 ACA 進行允入控制。但在執行時段(Execution Time Period)所出現之訊務並不規律，ACA 之剩餘資源也就不穩定(圖 4.11)，有可能因訊務超出預估值而造成資源不足，使得 ACA 無法再允入更多訊務，也可能因實際訊務低於預估值剩餘過多形成浪費。當 ACA 可掌握的資源過多，無法全數分配給進來的訊務，則會形成頻寬浪費，此時若是 BB 有頻寬回收的機制，則 ACA 可以繳回多餘的資源。當掌握的資源過少時，為了後續能順利允入新進來的訊務則必須以較高成本批購頻寬或以資源再分配的方式重新分配 Ingress Router 內部之頻寬。因此我們需要在執行時段設計一套執行時段頻寬管理機制，依照該時期內各個不同時間點與當時 ACA 所掌握的資源狀況來調整臨時批購決策，以維持適當頻寬存量以備允入後續出現之訊務。

執行時段頻寬管理機制的程序如下：在執行時段，ACA 對於 Ingress Router 所掌握之各 service class 會設定一個資源底限 y^* ，根據該時段的剩餘時間修正 y^* 值，以維持合理之資源存量，並於實際執行時段定時每隔一段時間檢查目前所擁有的資源是否大於底限，達到足以應付接下來可能出現之訊務，而時間間隔長度宜根據臨時補足頻寬所耗費的時間(On-demand order latency)訂定。

在執行時段頻寬管理機制下，於每個執行時段(Execution Time Period)根據 On-demand order latency 之時間長度設數個 check point，在每個 check point 檢查目前頻寬存量是否達到該 check point 所希望維持的資源底限 y^* 。隨著 check point 離執行時段結束時間越近，該 check point 之資源底限 y^*_c 也就越小，以避免保留資源過多使用不完形成浪費。

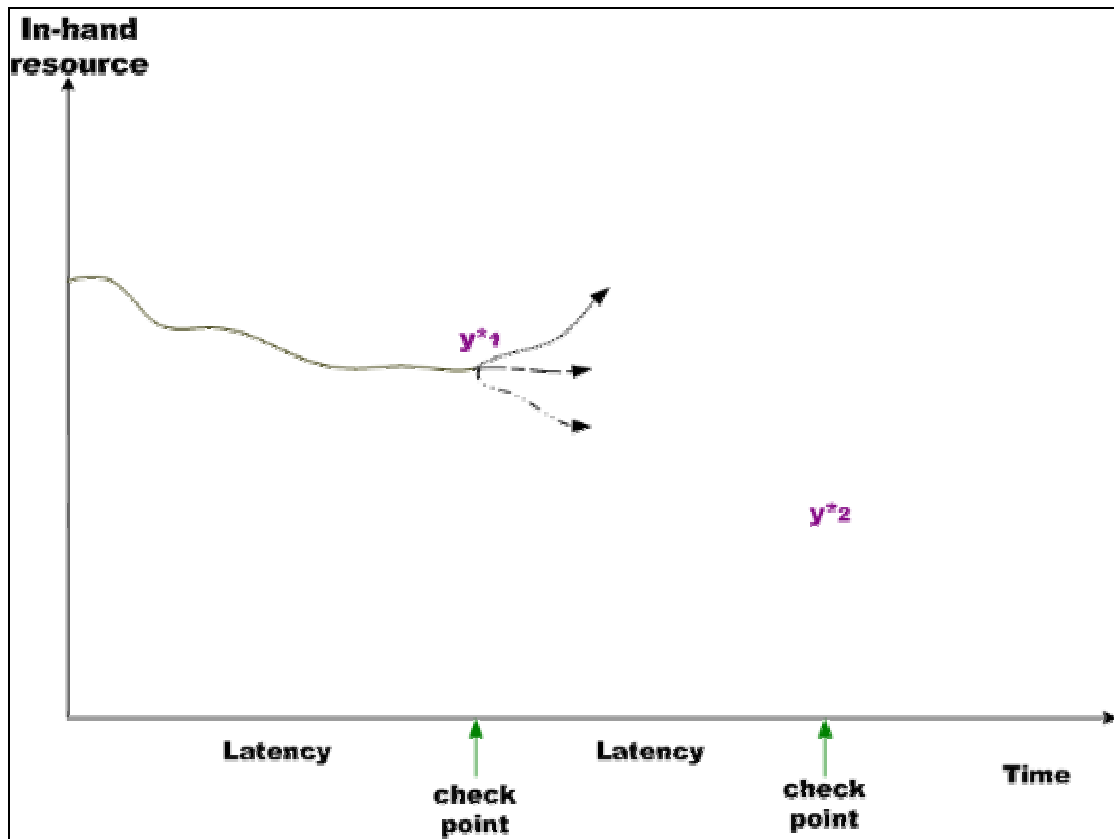


圖 4. 11: 處理未來之訊務需求所需資源水準

4.2.3 執行時段資源不足之解決方案

ACA 於執行時段發現資源不足時，可以根據 BB 於預購階段採用的配置策略之不同，選擇以資源再分配(resource reallocation)、臨時批購(on-demand request)或訊務降級(traffic flow degradation)等三種不同方式來解決資源不足的問題。

4.2.3.1.資源再分配

ACA 可先嘗試將所擁有多餘之資源重新分配。資源再分配是將在同一個 ACA 所擁有的資源，取相同服務等級互補，也就是不同路徑上具有相同的 quality entropy 之 link 頻寬可以互相流用。當一個新產生之訊務提出允入申請，ACA 根據目前可掌握的頻寬發現該訊務所要求之頻寬超出與目前可用路徑之頻寬，

ACA 可要求 Local SPPA 一一檢視該路徑(path)所經之 link,找出不足頻寬之 link,嘗試在這些 link 上作資源重分配,使用這些 link 的路徑如有超額存量,其配額即可考慮流用。而流用之決策尚待未來之研究深入探討。

4.2.3.2. 訊務降級

當新產生的訊務要求某一服務品質之路徑資源,但是 ACA 當時所掌握的資源數量不足,無法滿足該訊務時,若是該訊務可接受以較低等級之服務品質,則 ACA 可採用降級服務的方式,將該訊務所要求之服務品質降級,並檢視降級後相對應的服務等級資源是否足夠,若是可以滿足該訊務所提出的頻寬數量,則採用降級服務的方式允入該訊務。訊務降級雖然可以用替代的方式繼續維持允入服務,但是在各個服務等級的資源皆不充裕的情況下,降級服務的方式就再不適用。

4.2.3.3. 臨時批購

在執行時期,如果 BB 仍有保留統籌頻寬(中央保留法,或頻寬回收)以應付各 Ingress Router 資源不足時之臨時批購請求。於執行時段 ACA 直接根據目前最佳頻寬存量 y^*_c 與目前實際頻寬存量資源數目之差額透過 BOA 向 BB 發送臨時批購 message 進行即時的資源要求,以維持安全頻寬存量的方式預防臨時頻寬不足的現象。在每個 check point, ACA 會根據 check point 離執行時段結束之時間長度計算出所需維持的頻寬 y^* 。圖 4.12 展示執行時段頻寬管理流程。經過 BOA 向 BB 批購頻寬後,再由 PPA 重新計算形成可用之路徑資源,交由 ACA 對於後續之訊務進行允入控制。本研究以臨時批購的方式處理於執行時段維持安全存量以預防資源不足的問題,並於第四章以模擬的方式觀察效能與頻寬成本。

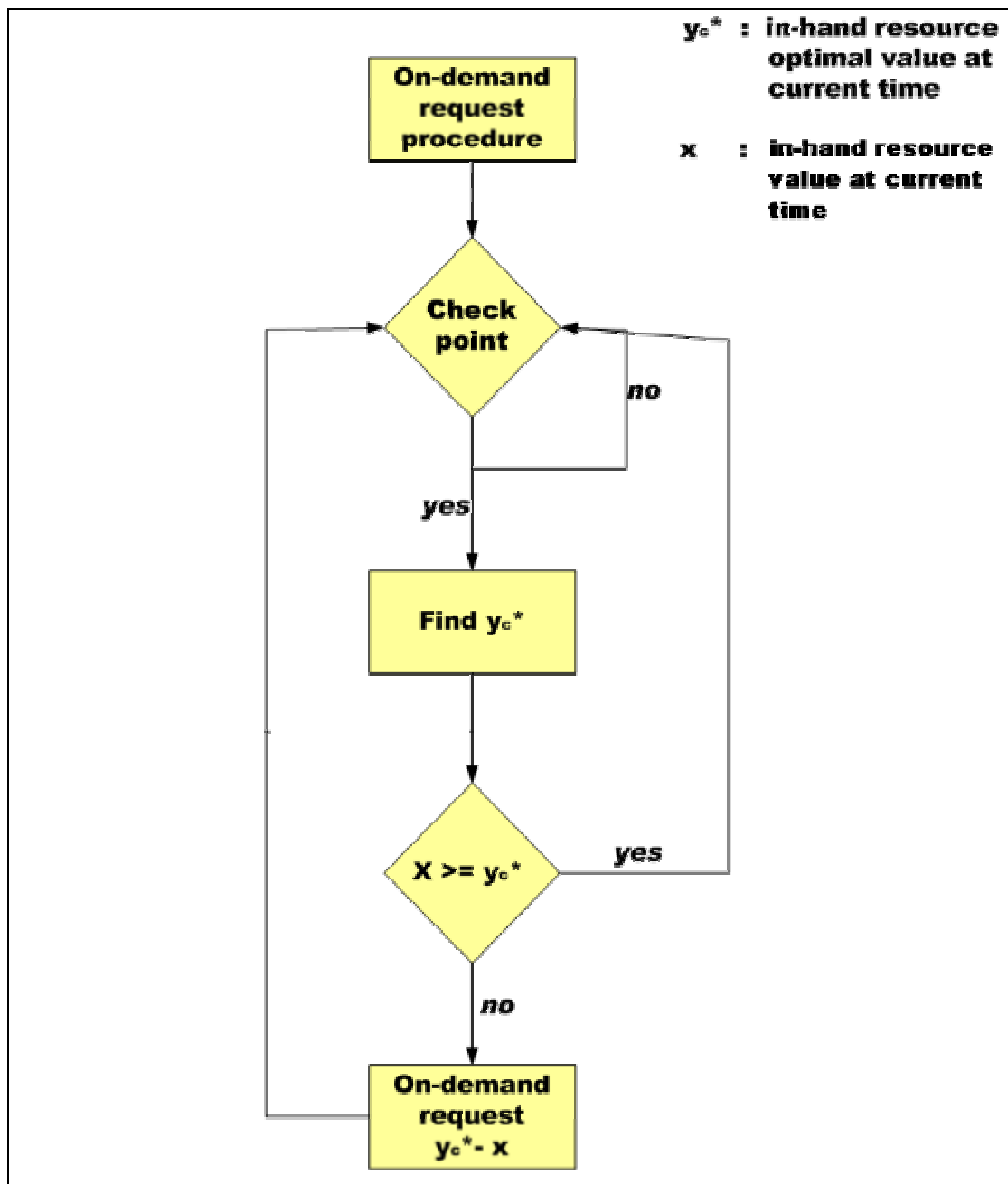


圖 4. 12: 執行時段頻寬管理流程

4.2.4 執行時段頻寬管理最佳化模型

在某個執行時段當中，ACA 選定幾個時間點(check point)檢查資源使用情形，以提供服務給可能出現之訊務。

我們以 $E\{B(x)\}$ 表示執行時段收益期望值(Execution Time Period expected profit)，藉由分析執行時段收益期望值，我們可以計算出各時段 ACA 之最佳頻

寬持有量。執行時段收益期望值分為兩部分，第一部分為臨時批購成本，其計算方法為預期頻寬持有量(y)與目前頻寬持有量(x)之差額乘上 Bandwidth Broker(BB)對於臨時批購索取的費率 C_2 。第二部分為 ACA 維持該時段所需之最佳頻寬對於接下來剩餘時段訊務分佈 $P_t(D)$ 之收益期望值(expected revenue)。於本架構中，假設 Ingress Router 對於每個允入之訊務可得到獲利 r ，若是拒絕允入訊務，則每拒絕一個訊務必須承受損失 p ，並且於同一個 Execution Time Period，單位獲利 r 與單位損失 p 之比例不會因為剩餘時間長短改變。

根據 Bandwidth Broker(BB)所訂定之每單位臨時批購價格(C_2)，各個 Ingress Router 對於允入每個訊務之獲利和拒絕允入之損失比例(r/p)，與剩餘時間之訊務分佈 $P_t(D)$ ，我們可以計算出該時間點之最大執行時段收益期望值，據此決定出該時間點之最佳頻寬持有量(y^*_t)。

Parameter List

- $P_t(D)$ PDF of traffic demand D at time t
- X Current in-hand resource
- Y Resource level
- y^*_t Optimal resource level at time t
- p Unit penalty for rejecting a request
- r Unit revenue for admitting a request
- C_2 On-demand unit cost
- $E\{B(x)\}$ Execution Time Period expected profit

每個 Execution Time Period 內不同時間點(check point)之 expected profit $E\{B(x)\}$ 為臨時批購成本與收益期望值之加總，如 Eq.4-5(a)表示。

$$E\{B(x)\} = -C_2 * (y - x) + \int_0^y rD * P_t(D)dD + \int_y^\infty [ry - p(D - y)] * P_t(D)dD \quad (4-5 a)$$

由於 Eq. 4-5(a)必須在所使用的頻寬機率分佈 $P_t(D)$ 為連續時才能夠使用，但

是關於頻寬之統計方式有可能為了方便起見，採用不連續的方式(discrete)作紀錄，所以我們將 Eq. 4-5(a)稍作修改成 Eq. 4-5(b)，以供不連續之頻寬記錄方式計算執行時段收益期望值。

$$E\{B(x)\} = -C_2 * (y - x) + \sum_{D=0}^y (rD * P_t(D)) + \sum_{D=y}^{\infty} ((ry - p(D - y)) * P_t(D)) \quad (4-5 b)$$

當目前持有頻寬為 x 時，若是 x 小於最佳頻寬 y ，則必須向 BB 批購 y 與 x 之差額 Eq. 4-6，服務往後會出現之訊務。

$$C_2 * (y - x) \quad (4-6)$$

對於最佳持有頻寬為 y 時，我們以進入的訊務與 y 的關係進行討論。假設進入訊務(incoming traffic)為 D ，發生的機率為 $P_t(D)dD$ ，當訊務需求(traffic demand) D 小於最佳頻寬持有量 y 時，收益為 $r * D$ 。對於最佳頻寬持有量 y 大於訊務需求 D 之期望值我們以 Eq.4-7(a)、Eq.4-7(b)表示之。

$$\int_0^y rD * P_t(D)dD \quad (4-7 a)$$

$$\sum_{D=0}^y rD * P_t(D) \quad (4-7 b)$$

當訊務需求 D 大於最佳頻寬持有量 y 時，收益為單位收益 r 乘上最佳頻寬持有量 y ，且不足的部分會乘上單位損失 p 作為因頻寬不足拒絕允入訊務之損失。對於最佳頻寬持有量 y 小於訊務需求 D 之期望值我們以 Eq.4-8(a)、Eq.4-8(b)表示之。

$$\int_y^{\infty} [ry - p(D - y)] * P_t(D)dD \quad (4-8 a)$$

$$\sum_{D=y}^{\infty} ((ry - p(D - y)) * P_t(D)) \quad (4-8 b)$$

本研究在某個執行時段內每個 check point，依其所在時間 t 修正最佳頻寬持有量 y ，並以 on-demand request 向 BB 要求頻寬補足不足之部分。根據時間點 t 不同與最佳頻寬持有量 y ，執行時段收益期望值也就不同。由於 check point 在本研究中時間是固定的，所以我們必須根據該時間點離執行時段結束時間之長度決定最佳頻寬持有量 y ，使得執行時段收益期望值之期望值為該時間的大值。

如果 Eq.4-5(a)中的 $P_t(D) \cdot D$ 與 $P_t(D)$ 皆有反導函數，最佳頻寬持有量可計算而得如下(Eq.4-9)。

$$\int_0^{y^*} P_t(D) dD = (p + r - c) / (p + r) \quad (4-9)$$

從 Eq.4-9 我們可以得知，假設 penalty，on-demand cost 與 revenue 皆不會在執行時段進行更動，由於根據 $P_t(D)$ 所代表的分佈維持頻寬持有水準。