Ripartiamo da MVA, ci sono le slides.

Abbiamo una rete a capacità finita.

Nel caso aperto, se non c'è spazio, lo porto via e lo conto nella probabilità di perdita.

Nell'immagine solo il centro 1 ha capacità finita, gli altri due no. Se questo diventa saturo, dovrebbe bloccare l'invio degli altri centri rispetto sè stesso.

Quindi dovremmo "fermarla/bloccarla", non toglierla. In questi contesti la regolarità si perde, si perde la struttura "a lattice". E' stato dimostrato che la *struttura a lattice è legata alle forme prodotto*, ovvero dove c'è tale struttura posso definire tale forma. Intituitivamente, se la forma è regolare, posso scrivere *equazioni simili* che la definiscono. La rete considerata *non è separabile*. E' stato proposto un algoritmo di MVA approssimato, chiamato **DEpenendent BEheviour MVA**, o **debeMVA**.

Vediamo il teorema degli arrivi $A_j(k) = E[n_j(k-1)]$, qui ne viene proposta una nuova forma in cui si aggiunge $z_j(k)+1$. Essendo capacità finita l'unica soluzione è sempre con il processo di Markov. I grafici sono indici di prestazioni locali nei centri.

Parte seconda di operAnOfQ

Abbiamo $T,\ A_i,\ B_i,\ C_{i,j}$ con $i=1,...,K$ con K numero di server (non più M). Sia $C_i = \sum_{j=0}^{K} C_{i,j}$, i completamenti ad i uscenti. Abbiamo anche gli arrivi dall'esterno: $A_0 = \sum_{j=1}^{K} A_{0,j}$ e $C_0 = \sum_{i=1}^{K} C_{i,0}$

Possiamo definire $U_i = \frac{B_i}{T}$, $S_i = \frac{B_i}{C_i}$ e $X_i = \frac{C_i}{T}$, ovvero utilizzazione, tempo di servizio e frequenza di uscita su i. Parliamo anche di probabilità di routing: $p_{i,j} = \frac{C_i,j}{Ci}$ se $i = 1, ..., k$, altrimenti se $i=0$ si ha $p_{i,j} = \frac{A_{0,j}}{A_0}$

Se sommiamo $p_{i,0} + \sum_{j=1}^{K} p_{i,j} = \frac{C_{i,0}}{C_i} + \sum_{j=1}^{K} \frac{C_{i,j}}{C_i} = \sum_{j=0}^{K} \frac{C_{i,j}}{C_i} = \frac{1}{C_i} \cdot C_i = 1$

Per definizione, la frequenza di uscita dell'intero sistema è $X_0 \doteq \frac{C_0}{T} = \sum_{i=1}^{K} \frac{C_{i,0}}{T} \cdot \frac{C_i}{C_i} = \sum_{i=1}^{K} \frac{C_i}{T} \cdot \frac{C_{i,0}}{C_i} = \sum_{i=1}^{K} X_i \cdot p_{i,0}$

$A_i \rightarrow ||||]O_{D_i} \rightarrow C_i \qquad n_i(t)$

$\bar{n_i} = E(n_i) = \frac{W_i}{T}$ ovvero sto calcolando area diviso il tempo. Vedendo il gradico "Queueing Networks" sulle y abbiamo proprio $n_i(t)$, e quindi sto cercando l'altezza media, cioè somma di quanto è alto per quanto tempo.

Il tempo di risposta medio per servire il singolo job è $R_i = \frac{W_i}{C_i}$ cioè il tempo per *singolo job* accumulato dalla risorsa i. W_i è tutto il tempo occupato dalla risorsa i durante l'osservazione, ma è di tutti i job, per questo dividendo per i completamenti, mi sto concentrando sul *singolo job*. Possiamo notare che $\bar{n_i} = \frac{W_i}{T} \cdot \frac{C_i}{C_i} = X_i \cdot R_i$ cioè **Little vale indipendentemente dalla condizione di job flow balance.** Little vale di più rispetto ai casi in cui è stato verificato. Supponiamo $A_i = 7\ j, B_i = 16\ s, C_i = 10\ j$, abbiamo: $n_i(0) = 3\ n_i(20) = n_i(0) + A_i - C_i = 0$ coerente con la figura! Qui ho considerato tempo iniziale e finale, se terminasse prima non avrei 0 ma Littel varrebbe comunque! $U_i = \frac{16}{20} = 0.8\ ,\ S_i = \frac{B_i}{C_i} = 1.6\ s\ , X_i = \frac{C_i}{T} = 0.5\ j/s\ \ , W_i = 40\ j \cdot s$

$\bar{n} = \frac{W_i}{T} = 2\ \ , R_i = \frac{W_i}{C_i} = 4\ s$

verifichiamo che $\bar{n_i} = X_i \cdot R_i = 0.5 \cdot 4 = 2$ Quindi **Little vale anche in assenza di job flow balance**. Per periodi di osservazioni lunghi possiamo comunque assumere job flow balance.

Scriviamo il numero di completamenti al centro j

$C_j = \sum_{i=0}^{K} C_{i,j}$ cioè tutti i completamenti da tutti gli i verso il singolo j

$\frac{C_j}{T} = \sum_{i=0}^{K} \frac{C_i}{T} \cdot p_{i,j}$ ovvero $X_j = \sum_{i=0}^{K} X_i \cdot p_{i,j}$ Se rete aperta X_i è noto, altrimenti se la rete è chiusa dobbiamo fissare delle equazioni di traffico.

Rapporti tra visite viste operazionalmente

Visit ratio $V_i = \frac{X_i}{X_0} = \frac{C_i}{T} \cdot \frac{T}{C_0} = \frac{C_i}{C_0}$, cioè quanta parte del flusso del sistema passa per i. Analiticamente era il numero medio di visite nel centro i.

Inoltre otteniamo la **legge del flusso forzato** $X_i = X_0 \cdot V_i$, cioè il flusso in qualsiasi parte del sistema V_i ci dà, se siamo in ipotesi job flow balacne, il throughput in quella parte del sistema.

Esempio

I job generano mediamente 5 richieste al disco, il throughput del disco è $X_{disk} = 10\ req/s$. Quale è il throughput del sistema? In assenza della legge sopracitata, sarebbe difficile dirlo, perchè abbiamo informazioni su un disco, non del sistema. Ma grazie alla legge del flusso forzato possiamo dire che è $X_0 = \frac{X_{disk}}{V_{disk}} = 10/5 = 2\ j/s$

Ribadiamo che Legge di utilizzazione e Little valgono anche senza bilanciamento del flusso. Se c'è, possiamo scrivere equazioni di bilanciamento del flusso come $X_j = \sum_{i=0}^{K} X_i \cdot p_{i,j}$

Ciò che si vede è che tutto ciò uscente dalla teoria delle code, Markov, processi stazionari è convalidato anche da sistemi che non soddisfano tali ipotesi.