Efficiency

$$\varepsilon(n) = \frac{T_{id}(n)}{T(n)}$$ ← ideal

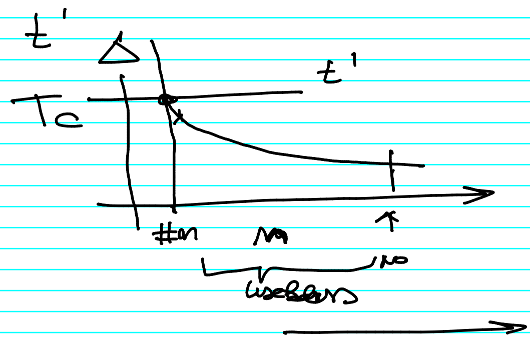
$$T_{id}(n) = \frac{T_{seq}}{n}$$

$$\varepsilon(n) = \frac{T_{seq}}{n\,T(n)} = \frac{sp(n)}{n}$$

speedup

good ε 85% ÷ 99%

$$\varepsilon(1) = 1 = \frac{T_{id}}{T(1) = T_{seq}} = \frac{T_{seq}}{T_{seq}} = 1$$



$$\varepsilon'(n) > \varepsilon''(n)$$

$$\frac{\cancel{T_{seq}}^{\,1}}{\cancel{n}\,T'(n)} > \frac{\cancel{T_{seq}}^{\,1}}{\cancel{n}\,T''(n)} \Rightarrow$$

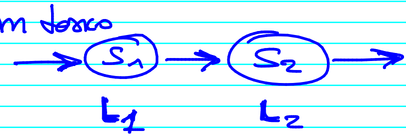
$$\frac{1}{T'(n)} > \frac{1}{T''(n)} \Rightarrow T'(n) < T''(n)$$

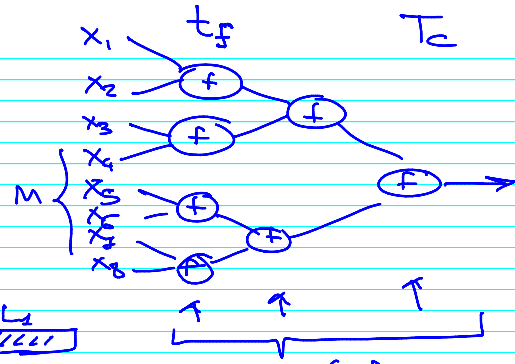
MOTIVATIONS

STRUCT. PROGR ENV

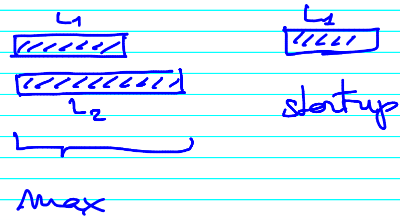
→ declared

m tasks





T_c

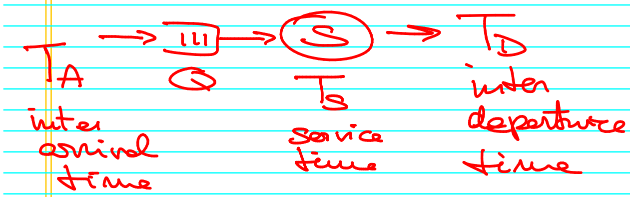


$T_c = \log_2(m)\, t_f$

$$T_c = L_1 + m\left(\max\{L_1, L_2\}\right)$$

QUEUE THEORY

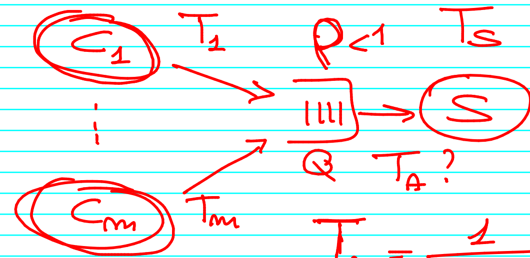
QUEUE THEORY



$$\rho_Q = \frac{T_S}{T_A}$$

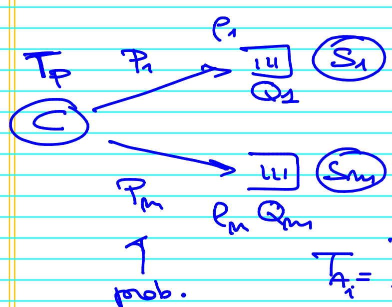
$\rho > 1 \Rightarrow$ queue grows, $T_D = T_S$

$\rho < 1 \Rightarrow T_D = T_A$



$$T_A = \frac{1}{\sum_1^m \frac{1}{T_i}}$$

$$T_A = \frac{1}{\sum_1^m \frac{1}{T_W}} = \frac{1}{\frac{m}{T_W}} = \frac{T_W}{m}$$



$$T_{A_i} = \frac{T_P}{P_i}$$

KIND & USAGE OF PERF MODELS

KIND $\left\{ \begin{array}{l} \text{SKELETON} \\ \text{TEMPLATE} \\ \text{ARCHITECTURE} \end{array} \right.$ MODELS

PIPELINE

$\to (S_1) \to (S_2) \to$

$\quad L_1 \qquad L_2$

$$T_S = \max\{T_{S_1}, T_{S_2}\} = \max\{L_1, L_2\}$$

↳ SKELETON MODEL

"better" implementation model

$T_{comm_{in}}$	$T_{comm_{in}}$
+	+
L_1	L_2
+	+
$T_{comm_{out}}$	$T_{comm_{out}}$

$$T_S = \max\{T_{comm_{in}} + L_2 + T_{comm_{out}}, T_{comm_{in}} + L_2 + T_{comm_{out}}\}$$

TEMPLATE MODEL

$$T_{comm}(d) = t_\emptyset + t_1 \cdot d$$

ARCHITE MODEL

dim of msg

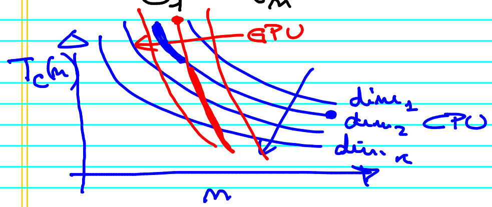
inverse of the network bandwith

COMPILE TIME USAGE

template choice

$T_1 \dots T_m$ $\qquad T_i \mid \min P_i$

$P_1 \quad P_m$

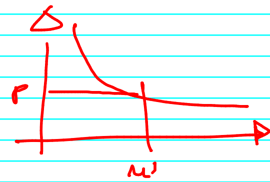


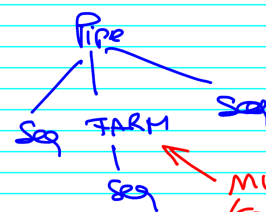
COMPARING ALTERNATIVES

CPU + GPU

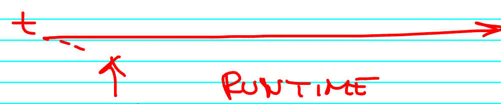
RUN TIME USAGE

$P(n) =$





MW = 5 (COMPILE TIME DECISION)

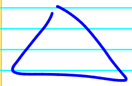


extra load

RUNTIME

1) OBSERVE P
2) COMPARE WITH THEORETIC ONE
3) ADAPT $\quad$ nw = 8?

POST RUN TIME USAGE



SK TREE $\Rightarrow$ tree of templates (perf models)

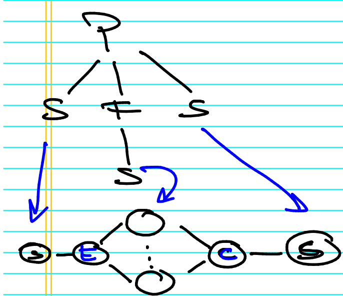
RUN PGM

\+ COLLECT MEASURES

BEFORE TERMINATION:

1) CONSULT THE P MODELS
2) OBSERVE DIFFS
3) TRY TO FIGURE OUT REASONS

$\rightarrow$ TELL the user (application programmer) : "PLEASE REMOVE FARM"



OBSERVE (BY MONITORING)

T_{S_1} T_E T_W T_C T_{S_3}

DIFF? $T_E \gg T_W$

normally $T_E = \frac{T_W}{nw}$

"CALCIUM" (PROACTIVE JAVA)

$\downarrow$

"SKANDIUM"

ACCURANCY OF PERF MODELS

VERY ACCURATE MODELS

APPROXIMATE MODELS

MUCH LONGER & MORE COMPLICATED FORMULAS WRT SIMPLER FORMULAS

EXPECTATION : MORE PRECISE MODELS

INTRODUCE COMPLEXITY

BEING ROOTED ON THE SAME PARAMS (SEQ TIME EXTIMATE)

PIPE

$$T_C \cong m \max\{T_{S_i}\}$$

#TASKS

FARM

$$T_C \cong T_W \left(\frac{m}{n_W}\right)$$