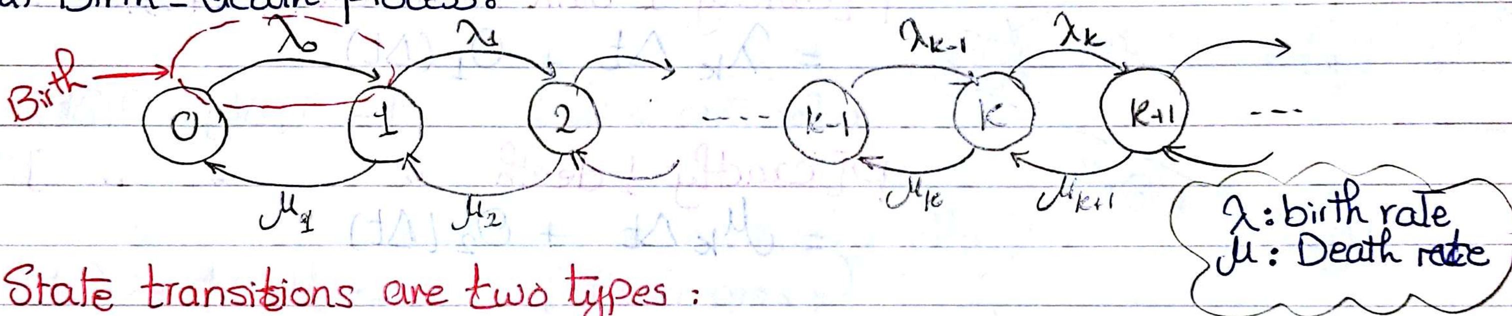
Unit (2) : Traffic Engineering

* ايه لازمة الـ traffic study ؟ عشان نعرف منها نعمل estimation كويس لأطوال الـ lines بتاعت الشبكة
وعشان نقلل الـ delay قدر المستطاع واعرف ايه الـ optimum buffer size اللي استخدمه .

(1) Queueing theory: mathematical study of waiting lines / queues
↳ Model is constructed so that queue lengths and waiting time can be predicted

⇐ بنستخدمها لما ناخد Business decision عن الـ Resources اللي محتاجينها عشان نـ provide Service ما.

(2) Birth-death Process:



State transitions are two types:
* Birth: increase the state variable by 1
* Death: decrease " " " " "

لـ الموديل دا بيستخدموه في حاجات كتير جداً؛ زي مثلاً الـ bacteria evolution أو مثلاً عدد الـ customers في line في السوبرماركت .

↳ Use (Birth-Death) in queuing theory:
* 3andna queue (FIFO), el queue da el arrivals feh bt-follow el Poission distribution (5dnah fy probability elsana elly fatet) el arrivals de (assume enha gaya mn ∞ population) and C Servers (with exponentially distributed Service time) w el queue has k places. (M/M/C/K/∞/FIFO)

* M/M/1 queue
↳ Single server queue with ∞ Buffer size, avg rate of arrival $= \lambda$
avg service time $1/\mu$. $\lambda_i = \lambda$ & $\mu_i = \mu$ for all i

System أن ال prob
يبقى عند state رقم k في زمن t :

Difference equation.

$$P'_k(t) = \lambda_{k-1} P_{k-1}(t) + \mu_{k+1} P_{k+1}(t) - (\lambda_k + \mu_k) P_k(t)$$

* M/M/C :

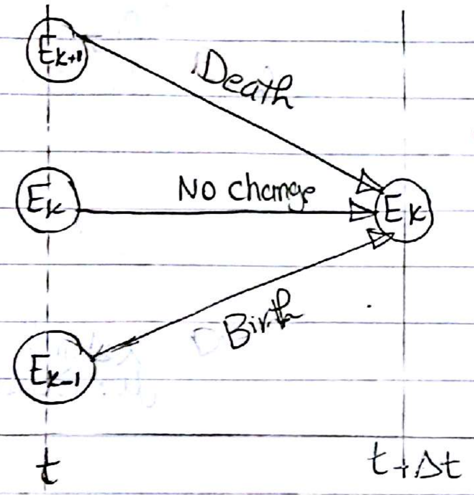
دا زي اللي فات بس الفرق بقى عندنا C servers وبالتالي ال Service time هيختلف

$\mu_i = i\mu \qquad (i \leq C)$

$\mu_i = C\mu \qquad (i \geq C) \qquad , \lambda_i = \lambda$ forall i

---

نرجع ال Slides بقى :D



P[exactly 1 birth in interval $(t, \Delta t + t)$ / k]

$= \lambda_k \Delta t + O_1(\Delta t)$

For Zero birth ? $= 1 - [\lambda_k \Delta t] + O(\Delta t)$

P[exactly 1 death " " "]

$= \mu_k \Delta t + O_2(\Delta t)$

Zero death ? $= 1 - \mu_k \Delta t + O(\Delta t)$

: probability of being @ state k at time $t + \Delta t$

$$P_k(t+\Delta t) = P_k(t) \underbrace{P_{k,k}(\Delta t)}_{\text{no change}} + P_{k-1}(t) P_{k-1,k}(\Delta t) + P_{k+1}(t) \underbrace{P_{k+1,k}(\Delta t)}_{\text{Death}} + O(\Delta t)$$

$k \geq 1$

دي معناها احتمالية اني

انتقل من state k-1 ل state k في زمن Δt اللي هو مثلاً من t_{n-1} ل t_n وليكن. (Birth)

← لو جينا نعمم المعادلات اللي فوق دي شوية ؟

$$P_k(t+\Delta t) = P_k(t) \underbrace{[1 - \lambda_k \Delta t + O(\Delta t)][1 - \mu_k \Delta t + O(\Delta t)]}_{\text{no change == 0 birth \& 0 death}} + P_{k-1}(t)$$

$$[\lambda_k \Delta t + O(\Delta t)] + P_{k+1}(t) [\mu_{k+1} \Delta t + O(\Delta t)] + O(\Delta t)$$

@ k=0

$$P_0(t+\Delta t) = P_0(t) [1 - \lambda_0 \Delta t + O(\Delta t)] + P_1(t) [\mu_1(t) + O(\Delta t)] + O(\Delta t)$$

and at k=0 ; $\mu_0 = 0 \Longrightarrow \sum_{k=0}^{\infty} P_k(t) = 1$ → #معروفة :D

Steady state

(3) Equilibrium: $P_k = \lim_{t\to\infty} P_k(t)$ — exist لو ال limit is in equilibrium بنقول ان ال queue

For M/M/1 queue →

$$\lambda_0 P_0(t) = \mu_1 P_1(t) \quad (k=0) \rightarrow ①$$

$$(\lambda_k+\mu_k) P_k(t) = \lambda_{k-1}P_{k-1}(t) + \mu_{k+1} P_{k+1}(t) \quad (k\geq 1) \rightarrow ②$$

flow rate into E_k = flow rate out of E_k

Assume:

$$\lambda_{-1} = \lambda_{-2} = \dots = 0, \quad \mu_0 = \mu_{-1} = \mu_{-2} = \dots = 0$$

$$P_{-1} = P_{-2} = \dots = 0$$

Equ 2 → solve at k=0

$$\lambda_0 P_0(t) + \underbrace{\mu_0 P_0(t)}_{0} = \underbrace{\lambda_{-1} P_{-1}(t)}_{0} + \mu_1 P_1(t) \longrightarrow \text{equ } 1 \checkmark\checkmark$$

$$\hookrightarrow P_1 = \frac{\lambda_0}{\mu_1} P_0$$

for k=1

* $\lambda_1 P_1(t) + \mu_1 P_1(t) = \lambda_0 P_0(t) + \mu_2 P_2(t)$

$$\hookrightarrow P_2 = \frac{\lambda_0 \lambda_1}{\mu_1 \mu_2} P_0$$

∴ Generally: $P_k = \dfrac{\lambda_0\lambda_1\lambda_2\dots\lambda_{k-1}}{\mu_1\mu_2\dots\mu_k} P_0$

$$\boxed{P_0 = \frac{1}{1+\sum_{k=1}^{\infty}\prod_{i=0}^{k-1}\frac{\lambda_i}{\mu_{i+1}}}}$$

---

* Notation used in Queue Models

a/b/c/d/e

- a → arrival process
- b → Service dis. or departure time
- c → No. of Servers
- d → No. of Customers
- e → Buffer space

لو آخر 2 دول مش معانا بنعتبرهم infinite.

Scanned by CamScanner

فا الـ M/M/1 اللى عاليه نقول عليه دا مثلاً

M / M / 1

- M → Markovain arraival process (poission)
- M → Markovain process (exponentially Distributed) service time & independent of arrival time
- 1 → 1 server

For M/M/1 : $\lambda_k = \lambda$, $k = 0, 1, 2, \dots$
$\mu_k = \mu$, $k = 1, 2, \dots$

$$\therefore P_k = P_0 \prod_{i=0}^{k-1} \left(\frac{\lambda}{\mu}\right) = P_0 \left(\frac{\lambda}{\mu}\right) \qquad k \geq 0$$

@ Equlibrium : $P_0 = 1 - \frac{\lambda}{\mu}$, Utilization factor $(\rho) = \frac{\lambda}{\mu}$

لـ الهري اللى فات دا كله ممكن نستخدمه في ايه مثلاً ؟ في الـ Network Router

↳ Hanshra7 elrasma elly fy slide 20

* الـ Router اللى ادامنا دا بيتكون من ٣ حاجات mainly

1- input ports : دى اللى الـ data packets بيتجمعوا من اكتر من user (المربعات الملونة دى)
وهنلاقى ان كل واحدة واقفة مستنية دورها علشان تخش على الـ Switching fabric
وهى واقفة مستنية في Buffer يعني طبعاً عندنا input Buffer (في queuing
اهه دحلوة :

2- Switching fabric : دا بقى على اساسه كل packet بتعرف هتطلع على انهى output port
ودا Based on algorithm معين مش مهم نعرفه المهم ان النتيجة

packet on certain port holding certain IP address → Algorithm → ytalla3ha 3ala outport no. x for ex.

٣- output ports : اللي قولنا هنطلّع عليها الـ packets ودي كمان ممكن يحصل عندها queue
برضو .. ليه؟ تخيل ان فيه port بطيء والـ switch fabric عمال
يطلع عليه packets! ساعتها حصل queue هنا.

→ Forwarding means ntala3 mn input port to output port

كدا module 1 خلص ، المهم فيه فهمنا ايه الـ Queuing th. والـ Birth_death وعرفنا
ازاي نوصف الـ queue والـ arrival فيه ، الـ Service time ، ومسائل على الكلام دا او بيتطبق
ازاي

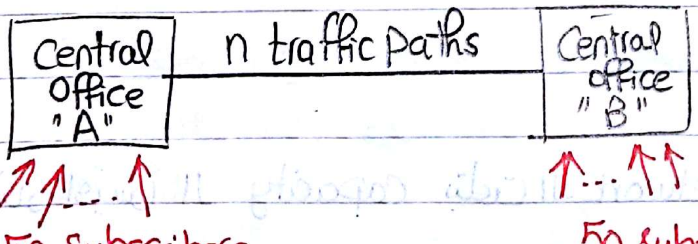
في Module 2 بقى عايزين نستخدم الـ traffic analysis كأنه شغل Quality كدا مثلاً

Module (2) : Traffic Charachristics & Measurements

* Performance analysis methods → Traffic engineering
* Steps :
  traffic Characterization (calling pattern) → System modeling (queue models, simulation) → analysis (blocking probab.)

* Given expected traffic + growth assumptions + ... , what's our design objectives?

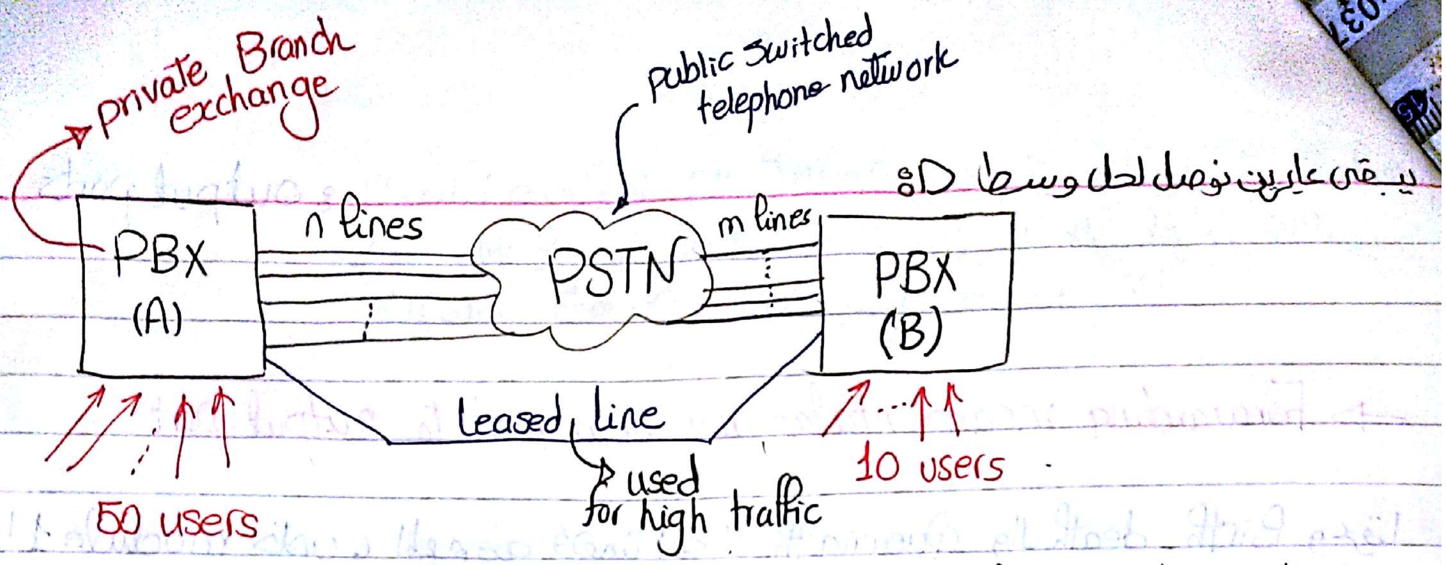
١- نحدد الـ resources اللي محتاجينها (switches, trunks, ...)
٢- نقلل الـ cost
٣- نوصل الـ minimum accepted Quality



* في الـ model ، لو وصلنا الـ 50 channels بـ 50 (n=50) دا افضل حاجة ومفيش Blocking
فكدا كله متوصل ببعضه اصلاً → High cost

* لو خلينا بقى n=1 فكدا الـ cost قلت
تماماً بس الـ Customers هيولعوا فينا :D
← كدا احتمالية الـ Blocking علت اوي

← Blocking يعني معرفتش انفذ الـ Call بتاعتي



PBX : is a telephone System within Enterprise that Switches Calls between enterprise users on local lines .

لو عندنا شركه فيها 2 offices وخليت

high cost, No block ← $m = 10$, $n = 50$

low cost, high block ← $m = 1$, $n = 1$

زى المثال اللى فات بس قولت مترمكوش من حاجه :D

---

* Traffic Characterization :

1- Call arrival → Random process

2- Call holding time → Random variable (الوقت اللى فيه مدته ال call traffic path بتاعنا)

3- Traffic intensity → msh sabta 3ala madar elyoum

→ Kasafet elcalls fy elpath, w bn-design based on Busy hour (Worst Case)

4- Busy hour : 1-hour period of the day during which traffic intensity is highest

Refers to traffic volume or number of calls attempts

---

* Traffic measurments :

* عشان اقيس ال capacity بتاعت ال network يبقى هنقيس ال Volume of traffic ال فى فترة زمنية معينة

→ Traffic intensity = $\frac{\text{traffic Volume}}{\text{Length of time}}$ → elly ba2es feh

used to measure avg. utilization during time period

How to Measure telephone traffic?

(1) Traffic intensity (A): $A = \lambda t_m$

λ → avg arrival rate (Calls/hr)

t_m → avg holding time in hrs

Units → Erlangs (dimensionless): Calls - Second per second

CCS (Calls-second per hour); 1 erlang = 36 CCS

For Ex: Capacity of single channel is 1 erlang

telephone busy 10% of time → load of 0.1 erlang on that line

→ $\boxed{\text{2 Calls/hour}}$ (λ) with avg holding time of 5 min. → t_m

A in Eralngs: $\frac{2 \times 5}{60} = 1/6$ Erlang

A in CCS: $\frac{1}{6} \times 36 = 6$ CCS

1 eal. → 36 ccs

1/6 → ?

كدا عرفنا ازاى بنقيس traffic ال channel وال load ل given ال avg arrival rate وال hold time

عاوزين نتكلم عن ال arrivals بقى

---

* Arrival Distribution:

Most fundamental assumption is that "Call arrivals are independent"

← فنزة ال math بقى :D

السنة اللى فاتت اخدنا ال Poission distribution فى ال Discrete ← time ~~Continuous~~

الموضوع ايه بقى ، عندنا شوية independent sorces عمال بيجيلنا منهم arrivals

probability ↳ $P(x)$ ← arrivals

$$P(x) = \begin{cases} \dfrac{e^{-\alpha}\alpha^{x}}{x!}, & x = 0, 1, 2, \ldots \\ 0, & \text{other.} \end{cases}$$

In telephone network, we assume "independent arrivals (calls)" with arrival rate λ

for $A = \lambda t_m$

$\therefore P(x \text{ arrivals during time } t_m) = e^{-A} \dfrac{A^x}{x!}$

↳ Prob. of one or more arrival $= 1 - \sum_{i=0}^{j-1} P_i(A)$

نحل مثالين على الكلام دا بقى

Slides (33,34)

Ex(1): node with $\lambda = 4$ msg/min, find probability of 8 or more arrivals in 30 sec.

$\tau = 30$ Sec , $\lambda = \frac{4}{60} = 1/15$

$A = \lambda\tau = \frac{1}{15} * 30 = 2$

$$P(j \geq 8) = 1 - \sum_{i=0}^{7} P_i(\lambda\tau)$$

$= 1 - [P_0(2) + P_1(2) + \dots + P_7(2)]$

$= 1 - \sum_{i=0}^{7} \left(e^{-2} \frac{2^i}{i!}\right) = 1.09 * 10^{-3} \simeq 0.00109$

Ex(2): enough channels, $\lambda = 1$ call/min, $\tau_m = 2$ min

What percentage of total traffic is carried by first 5 Circuits?

How much traffic carried by remaining Circuits?

$A = \lambda \tau_m = 1 \times 2 = 2$ erlangs

→ Traffic intensity carried by 1 active device is 1 erlang

$\therefore$ 2 erlangs ⟶ 2 devices

For total 5 :

$A_5 = 1 * P_1(2) + 2 * P_2(2) + 3 * P_3(2) + 4 * P_4(2) + 5 * P_5(2)$ لوحد عارفين منيه بيقولك ٧:

* Remaining Circuits will carry = $2 - A_5$

Holding time:

2 Common assumptions → constant holding time / → exponentially distributed

→ Constant: Suitable for call setup, interoffice signaling, recorded msg playback.

→ Exponential Distribution:

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & \text{else} \end{cases}$$

Mean = $1/\lambda$, variance = $1/\lambda^2$

= Used to model interarrival times (random) and service times that are highly variable.

* probability of a call termination independent of how long a call has been in progess.

$$P(\tau > t) = e^{-t/t_m}$$

→ probability that holding time exceed "t"

→ Negative Exponential:

- Arrivals are indepenelent

1) in small interval Δt only one arrive can occur
2) Probability of arrival $\propto \Delta t$ وده طبيعي، كل ما المدة هتزيد احتمالية اني يجي call بتزيد
   → & $= \lambda \Delta t$
   → avg call arrival rate from large group of independent sources
3) ال arrivals في interval معينة ملهاش علاقة بال arrivals في interval تانية
4) time between 2 arrivals (exponential) → $P_0(\lambda\tau) = e^{-\lambda\tau}$

Ex (Slide 42)

10.000 Subscriber lines originate one call per hour, how often two calls arrive with less than 0.01 sec between them?

0.01 sec

Call 1 Call 2

↳ time between 2 arrivals is exponentialy distributed

$$P_0(\lambda\tau) = e^{-\lambda\tau}$$

$$\lambda = \frac{10000}{60 * 60} = 2.78 \text{ arrivals/sec}$$ (10000 by3mlo mokalma kol sa3a ya3ny each hr 10000)

* probability of arrivals in 0.01 s

$$P_0(\lambda\tau) = e^{-2.78 * 0.01} = 0.973$$

Probability of an arrival $= 1 - P(\lambda\tau) = 0.027$

→ rate of occurance in 0.01 sec

$= \lambda *$ Prob. of occurence

$= 2.78 * 0.027$

Scanned by CamScanner