

Dimensionnement

Marceau Coupechoux
Département Informatique et Réseaux

Objectifs du cours

- Comprendre la notion de trafic
- Maîtriser l'application des lois d'Erlang dans un contexte de téléphonie

2

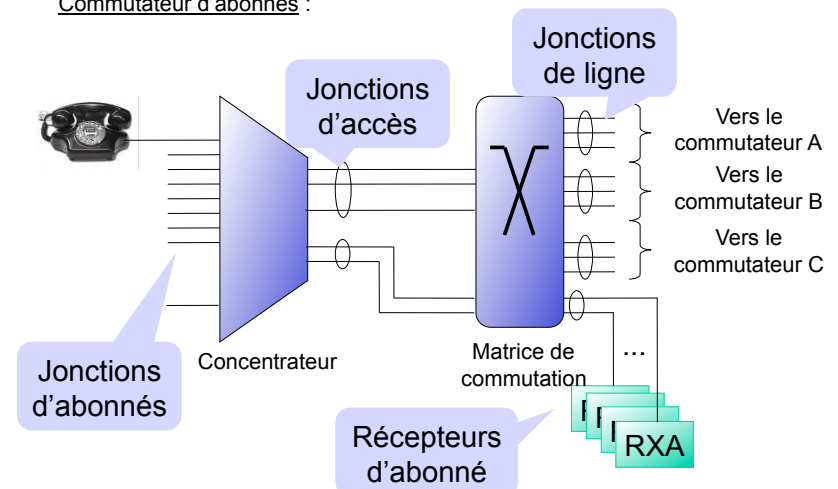
Sommaire

- Quelques problématiques de dimensionnement
- Notion de trafic
- Modèles à perte : Erlang-B, Engset
- Modèles à attente : Erlang-C
- La problématique du débordement
- Conclusion

3

I. Quelques problématiques de dimensionnement

Commutateur d'abonnés :



4

I. Quelques problématiques de dimensionnement

Objectifs du dimensionnement :

- Combien de RXA ?
- Combien de jonctions d'accès ?
- Combien de jonctions de ligne ?
- ... pour quelle qualité de service ?

Critères de qualité de service :

- Processus avec perte : taux de blocage, e.g. 1%,
- Processus avec attente : délai d'attente et quantile, e.g. moins d'une seconde dans 95% des cas.

Compromis investissement / qualité de service :

- Une meilleure qualité de service nécessite plus de jonctions, donc plus d'investissements.

II. Notion de trafic

Remarques :

- Le trafic moyen à l'instant u est :

$$a(u) = \lim_{T \rightarrow \infty} a(u, T)$$

- Pour un trafic stationnaire, cette quantité ne dépend pas de u .
- Le trafic s'exprime en **Erlangs**.

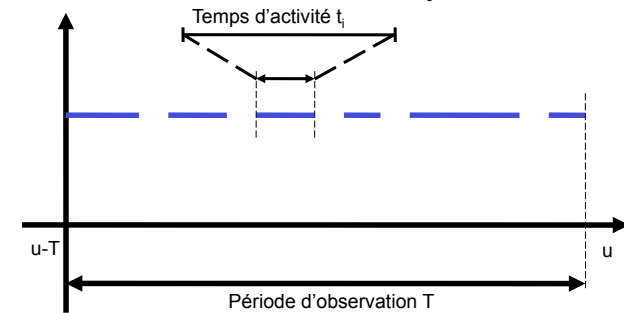
II. Notion de trafic

Trafic d'un serveur : c'est le temps moyen d'occupation du serveur, c'est aussi la probabilité d'occupation du serveur.

Sur une période d'observation T , le trafic vaut à u :

NB : pour un serveur $a \leq 1$

$$a(u, T) = \frac{1}{T} \sum_i t_i$$



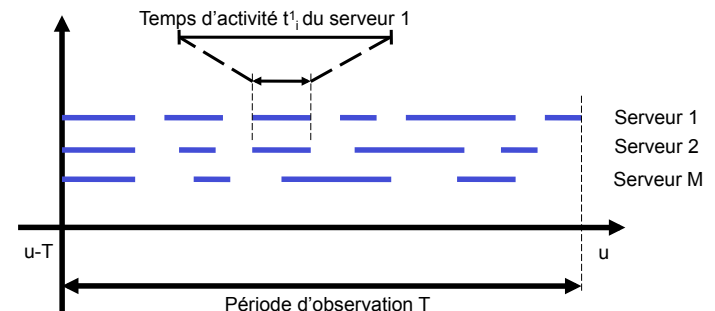
II. Notion de trafic

Trafic d'un groupe de serveurs : c'est la somme des trafics de chaque serveur.

Pour une période d'observation T , le trafic à u vaut :

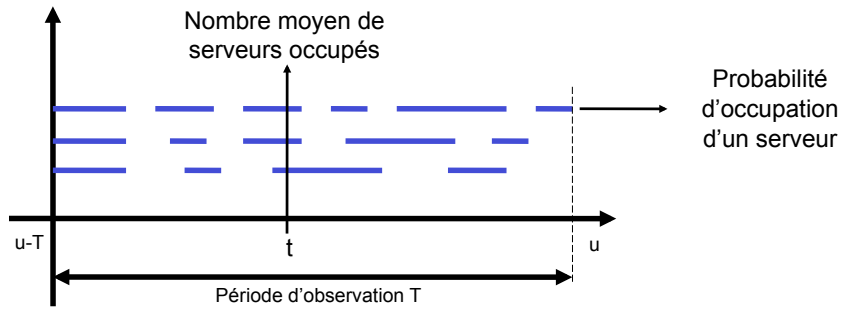
NB : pour M serveurs $A \leq M$

$$A(u, T) = \frac{1}{T} \sum_{j=1}^M \sum_i t_i^j$$



II. Notion de trafic

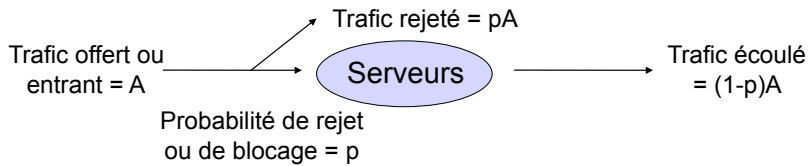
Trafic ergodique : le nombre moyen de serveurs occupés à un instant t est égal à la probabilité d'occupation d'un serveur.



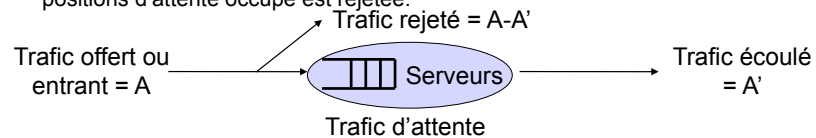
9

II. Notion de trafic

Processus à perte : une demande de service trouvant l'ensemble des serveurs occupé est rejetée.



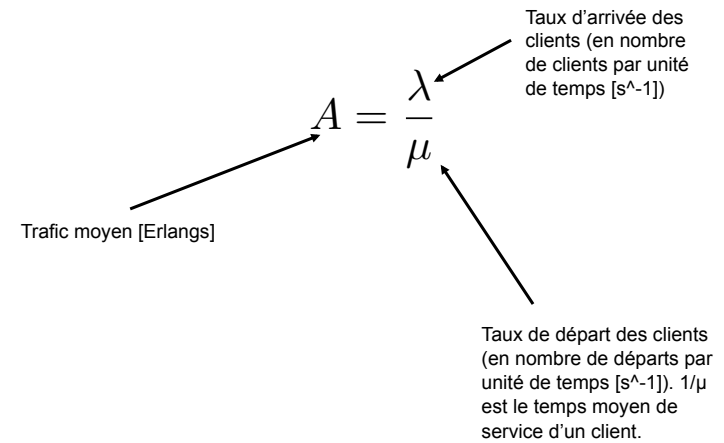
Processus à attente : une demande de service trouvant l'ensemble des serveurs occupé est mise en attente. Une demande de service trouvant l'ensemble des positions d'attente occupé est rejetée.



11

II. Notion de trafic

Autre formulation du trafic :



10

II. Notion de trafic

L'unité de trafic :

Le trafic est sans unité.

« L'unité » utilisée pour mesurer le trafic est l'**Erlang**.

Un serveur occupé à 100% du temps supporte un trafic de 1 Erlang.

Un groupe de M serveurs occupés à 100% supporte un trafic de M Erlangs.

12

III. Modèles à perte

I. Modélisation file d'attente



La notion de **trafic** en téléphonie recouvre la même notion que la **charge** en théorie des files d'attente.

Le trafic offert est caractérisé par le processus d'arrivée des clients. En téléphonie, le processus d'arrivée est **Poisson**, i.e. un processus de comptage N stationnaire à incréments indépendants tel que :

$$\Pr(N(s+t)-N(s) = k) = (\lambda t)^k e^{-\lambda t} / k!, \text{ pour tous } s \text{ et } t \text{ dans } \mathbb{R} \text{ et } k \text{ dans } \mathbb{N}.$$

Un processus de Poisson est caractérisé par le **taux d'arrivée** des appels/clients : λ appels/s. Les inter-arrivées suivent une loi exponentielle de paramètre λ .

La durée d'un appel suit un **loi exponentielle** de paramètre μ . Si T est le temps de service et $f_T(t)$ sa pdf :

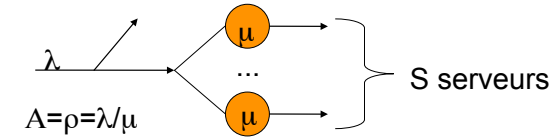
$$f_T(t) = \mu e^{-\mu t}$$

$$E[T] = 1/\mu$$

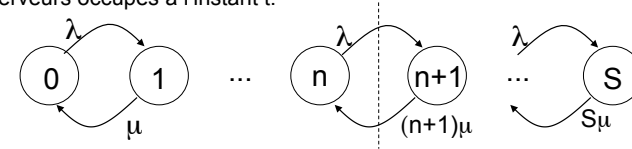
13

III. Modèles à perte

I. Modélisation file d'attente



Chaîne de Markov à temps continue : $X(t) = \{n(t)\}_{t \geq 0}$ où $n(t)$ est le nombre de serveurs occupés à l'instant t .



Equations aux frontières : $\lambda p(n) = (n+1)\mu p(n+1)$ pour tous n entre 0 et $S-1$.

$$p(n) = A^n / n! p(0) \text{ pour } n \text{ entre } 0 \text{ et } S.$$

Equation de normalisation : $\sum_n p(n) = 1$.

14

III. Modèles à perte

II. Loi d'Erlang B



Les probabilités stationnaires s'écrivent :

$$p(n) = A^n / n! / \sum_{0 \leq i \leq S} A^i / i! \text{ pour } n \text{ entre } 0 \text{ et } S.$$

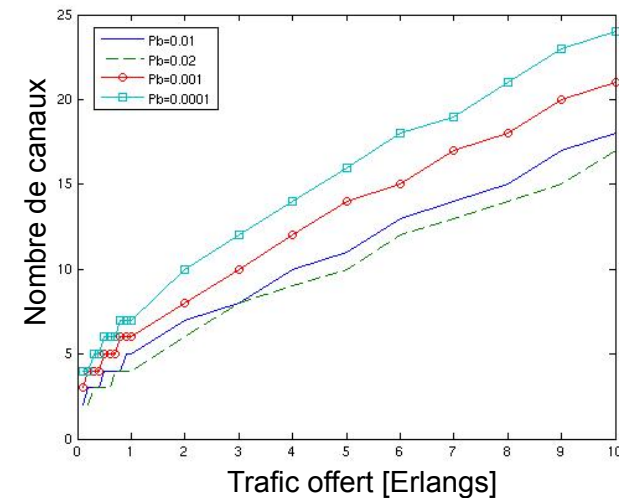
La probabilité de perte ou de blocage :

$$E_{1,S}(A) = \frac{\frac{A^S}{S!}}{\sum_{i=0}^S \frac{A^i}{i!}}$$

15

III. Modèles à perte

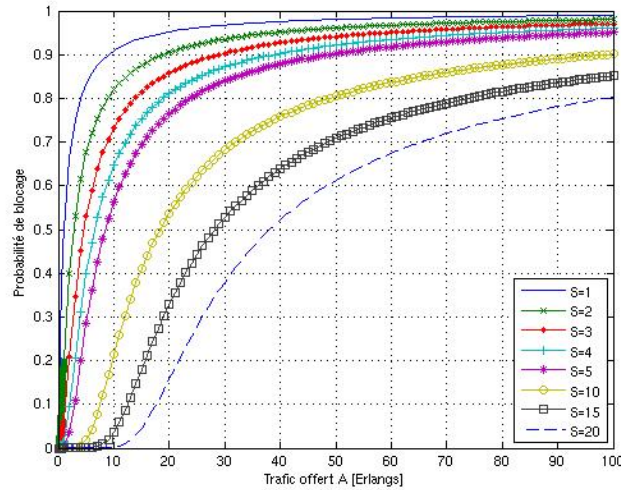
II. Loi d'Erlang B



16

III. Modèles à perte

II. Loi d'Erlang B

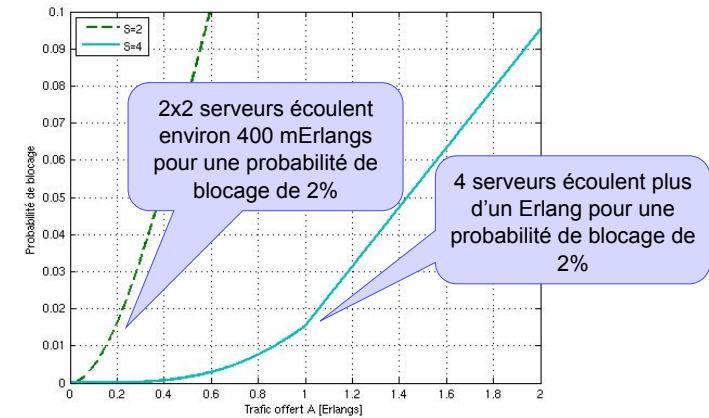


17

III. Modèles à perte

II. Loi d'Erlang B

Gain statistique (*trunk gain*) :



18

III. Modèles à perte

II. Loi d'Erlang B

Formule de récurrence permettant d'obtenir, pour un trafic offert donné, les valeurs successives du taux de perte en fonction des ressources S :

$$E_{1,S+1}(A) = \frac{AE_{1,S}(A)}{S + 1 + AE_{1,S}(A)}$$

Loi de Rigault [1] ou approximation de l'inverse de la loi d'Erlang B :

$$\text{Si } E_{1,S}(A) = 10^{-k}, \text{ alors } S \approx A + k\sqrt{A}$$

Exemple : 10 appels par min, appel de durée moyenne 2 min, probabilité de blocage de 1% > 30 jonctions (Rigault donne 29).

19

III. Modèles à perte

II. Loi d'Engset

On a supposé jusqu'ici un nombre infini de source et un taux d'arrivée des appels global pour l'ensemble des sources. L'hypothèse est réaliste pour un « grand » nombre de sources.

Si la population est finie, le taux des arrivées diminue avec le nombre d'appels en cours de traitement. A la limite, si toutes les sources sont servies, il n'y a plus de nouvelles arrivées.

Hypothèses :

- S serveurs.
- M > S sources Poisson indépendantes avec taux d'arrivée α .
- Temps de service exponentiel de paramètre μ .
- Trafic créé par une source : $a = \alpha/\mu$.

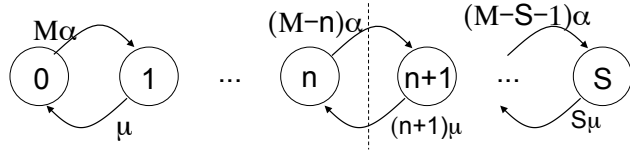
20

III. Modèles à perte

II. Loi d'Engset



Le taux d'arrivée dépend du nombre de sources en cours de service.



Equations aux frontières : $(M-n)\alpha p(n) = (n+1)\mu p(n+1)$ pour n entre 0 et $S-1$.

$p(n) = C_M^n a^n p(0)$ pour n entre 0 et S .

Equation de normalisation : $\sum_n p(n) = 1$.

Les **probabilités stationnaires** s'écrivent :

$$p(n) = C_M^n a^n / \sum_{0 \leq i \leq S} C_M^i a^i \text{ pour } n \text{ entre } 0 \text{ et } S.$$

III. Modèles à perte

II. Loi d'Engset



Probabilité d'engorgement : $p(S) = C_M^S a^S / \sum_{0 \leq i \leq S} C_M^i a^i$

Probabilité de blocage : $P_b = \alpha(M-S)p(S) / \sum_{0 \leq i \leq S} \alpha(M-i)p(i)$

$$P_b(M, S, a) = \frac{C_{M-1}^S a^S}{\sum_{i=0}^S C_{M-1}^i a^i}$$

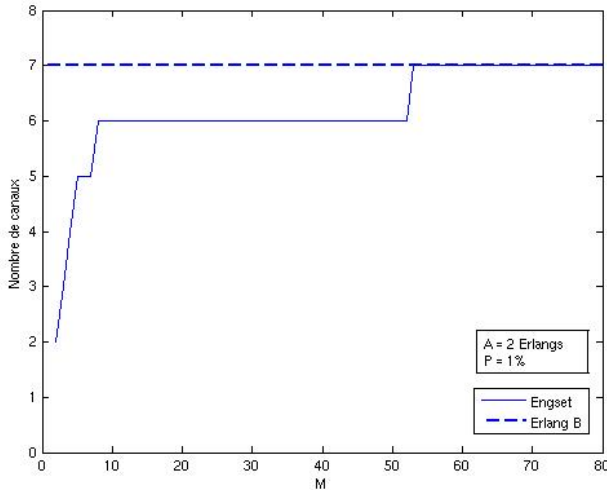
En pratique, lorsque $M > 10 \cdot S$, on utilise Erlang B qui donne un majorant très proche sur le nombre de jonctions à installer.

Remarque : le trafic moyen total s'exprime en fonction de a et P_b .

$$\begin{aligned} A &= \sum_{i=0}^S a(M-i)p(i) \\ &= \frac{aM}{1 + a(1 - P_b)} \end{aligned}$$

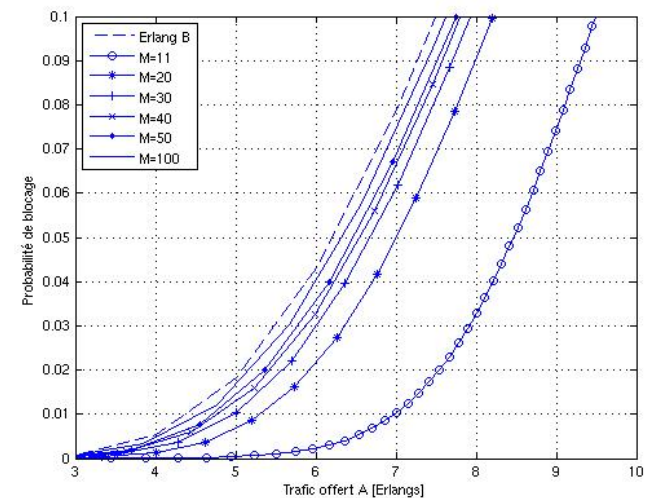
III. Modèles à perte

II. Loi d'Engset



III. Modèles à perte

II. Loi d'Engset



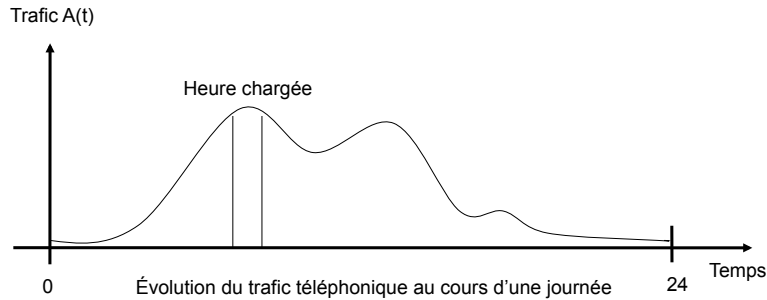
III. Modèles à perte

III. Heure chargée (*Busy Hour*)



En toute rigueur, le trafic téléphonique est une grandeur qui varie au cours du temps $A(t)$

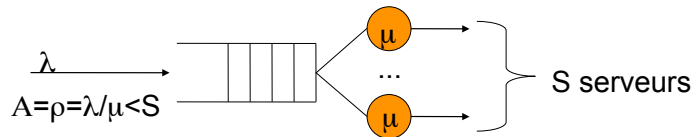
Pour dimensionner un système téléphonique, on prendra systématiquement le trafic à l'heure chargée comme référence



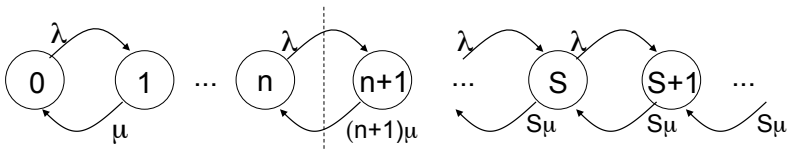
25

IV. Modèles à attente

II. Modélisation file d'attente



Chaîne de Markov à temps continue : $X(t) = \{n(t)\}_{n \geq 0}$ où $n(t)$ est le nombre de serveurs occupés à l'instant t .



27

IV. Modèles à attente

I. Applications de la loi d'Erlang C



Hypothèses :

- L'application de cette loi suppose que le nombre de clients est suffisamment important pour être considéré comme infini. Dans ces conditions le processus de service est indépendant du processus d'arrivée.
- Le processus d'arrivée des appels est Poisson de paramètre λ .
- Le temps de service suit une loi exponentielle de paramètre μ .
- La file d'attente est infinie (pas de pertes) et la politique de service est FIFO.

Application :

La loi C d'Erlang est utilisée pour dimensionner le nombre de récepteurs d'abonnés au niveau des centraux téléphoniques

Loi à attente => Le paramètre de performance important est la statistique du délai d'attente d'un client avant d'être mis en relation avec un récepteur d'abonné (processus de présélection).

26

IV. Modèles à attente

III. Loi d'Erlang C



Equations aux frontières :

$$\lambda p(n) = (n+1)\mu p(n+1) \text{ pour } n \text{ entre } 0 \text{ et } S-1,$$

$$\lambda p(n) = S\mu p(n+1) \text{ pour } n > S-1.$$

$$p(n) = A^n / n! p(0) \text{ pour } n \text{ entre } 0 \text{ et } S-1,$$

$$p(n) = A^n / (S! S^{n-S}) p(0) \text{ pour } n > S-1.$$

Equation de normalisation :

$$1/p(0) = \sum_{n=0}^{S-1} A^n / n! + A^S / S! \times 1/(1-A/S)$$

Loi d'Erlang C : probabilité d'attente

$$E_{2,S}(A) = \frac{\frac{A^S / S!}{1 - A/S}}{\sum_{i=0}^{S-1} \frac{A^i}{i!} + \frac{A^S / S!}{1 - A/S}}$$

28

IV. Modèles à attente

III. Loi d'Erlang C



Remarque : la loi d'Erlang C peut s'exprimer en fonction de la loi d'Erlang B.

$$E_{2,S}(A) = E_{1,S}(A) \{1 - A/S(1 - E_{1,S}(A))\}^{-1}$$

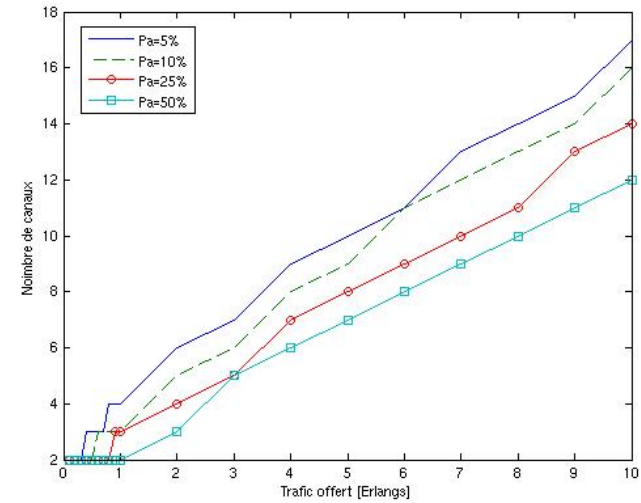
$$E_{2,S}(A) \approx E_{1,S}(A) / (1 - A/S)$$

Exemple : 10 appels par minute, durée moyenne d'appel de 2 min, probabilité d'attente maximale de 10%
> 27 jonctions sont nécessaires.

29

IV. Modèles à attente

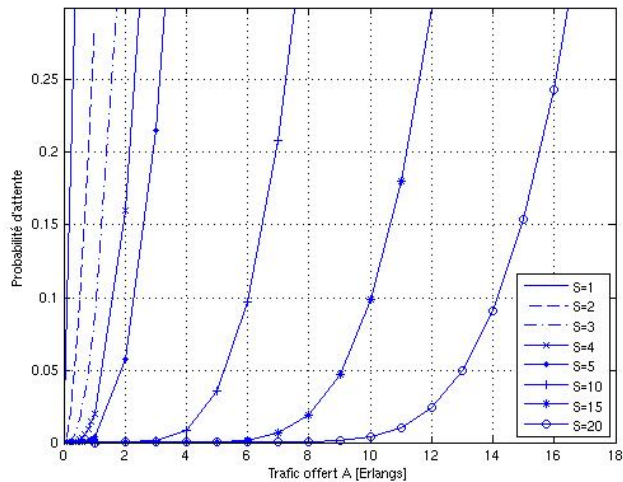
III. Loi d'Erlang C



30

IV. Modèles à attente

III. Loi d'Erlang C



31

IV. Modèles à attente

III. Loi d'Erlang C



Quelques résultats utiles : (cf. en annexe pour les justifications)

Nombre moyen de clients dans la file : $Q = E_{2,S}(A) \frac{A/S}{1 - A/S}$

Temps d'attente moyen dans la file : $W = \frac{E_{2,S}(A)}{\lambda} \frac{A/S}{1 - A/S}$

Distribution du temps d'attente dans la file :

$$P[T > t] = E_{2,S}(A) \exp(-S\mu t(1 - A/S))$$

32

IV. Modèles à attente

III. Loi d'Erlang C

Quelques résultats utiles : cas mono-serveur M/M/1

Nombre moyen de clients dans la file : $Q = \frac{A}{1 - A}$

Débit : $X = \lambda$

Taux d'utilisation du serveur : $U = A$

Temps d'attente moyen dans la file : $R = Q/X = \frac{1}{\mu - \lambda}$

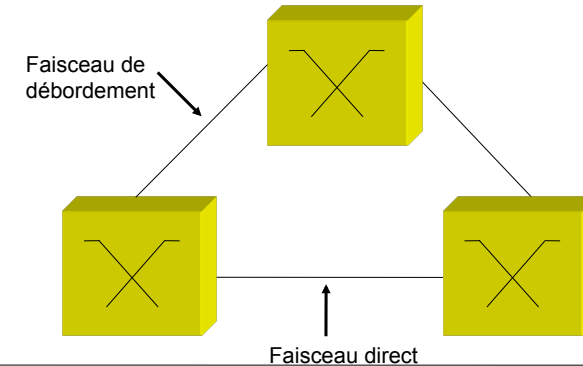
Distribution du temps d'attente dans la file : $P[T > t] = A \exp(-t(\mu - \lambda))$

V. La problématique du débordement

Le trafic rejeté par un premier système est récupéré et offert à second système.

Si le second système n'a plus lui-même de ressources, le trafic pourra être offert à un troisième système.

Il est aussi possible de regrouper sur un même faisceau l'ensemble des trafics de débordement venant de différents systèmes.



V. La problématique du débordement

Deux méthodes peuvent être utilisées pour dimensionner les faisceaux de débordement :

- ✓ La méthode du faisceau équivalent de Wilkinson
- ✓ Les processus MMPP

Conclusion

- Dimensionnement :
 - Raccordement à un récepteur d'abonnés : Erlang C
 - Dimensionnement des jonctions d'accès : Engset ou Erlang B
 - Dimensionnement des jonctions de ligne : Engset ou Erlang B
- Critères de qualité de service :
 - Systèmes avec perte : probabilité de blocage,
 - Systèmes à attente : probabilité d'attente ou délai/quantile.
- Erlang B et C sont aussi utilisées dans les réseaux de paquets et les réseaux mobiles.
- Pour approfondir : problématique du débordement (Wilkinson ou MMPP), réseaux multi-services, réseaux avec priorité, clients impatientes, répétition des appels infructueux, mixte de trafic (communications efficaces, non-réponses, occupation, numérotation incorrecte,...), etc.

Annexe : paramètres de performance M/M/K



$$\begin{aligned}
 \text{Nombre moyens de clients dans la file : } Q &= \sum_{n \geq S+1} (n - S)p(n) \\
 &= \sum_{n \geq 0} np(n + S) \\
 &= \sum_{n \geq 0} np(0) \frac{A^{n+S}}{S!S^n} \\
 &= \frac{p(0)A^{S+1}}{S!S} \frac{1}{(1 - A/S)^2} \\
 &= E_{2,S}(A) \frac{A/S}{1 - A/S}
 \end{aligned}$$

Temps moyen passé dans la file : Little appliquée à la file d'attente

$$\begin{aligned}
 W &= Q/\lambda \\
 &= \frac{E_{2,S}}{\lambda} \frac{A/S}{1 - A/S}
 \end{aligned}$$

37

Annexe : paramètres de performance M/M/K



- Le temps d'attente d'un client trouvant j clients dans la file est une variable aléatoire somme de $(j+1)$ variables exponentielles de paramètre $S\mu$: sa loi est Erlang- j (on le montre par récurrence) :

$$P[T > t | T > 0, Q = j] = \sum_{k=0}^j \frac{(S\mu t)^k}{k!} e^{-S\mu t}$$

- La cdf de T s'écrit alors :

$$\begin{aligned}
 P[T > t] &= \sum_{j \geq 0} \sum_{k=0}^j \frac{(S\mu t)^k}{k!} e^{-S\mu t} \frac{A^{j+S}}{S!S^j} p(0) \\
 &= \frac{e^{-S\mu t} A^S p(0)}{S!} \sum_{k \geq 0} \sum_{j \geq k} \frac{(S\mu t)^k}{k!} \frac{A^j}{S^j} \\
 &= \frac{e^{-S\mu t} A^S p(0)}{S!(1 - A/S)} \sum_{k \geq 0} \frac{(S\mu t A/S)^k}{k!} \\
 &= E_{2,S}(A) \exp(-S\mu t(1 - A/S))
 \end{aligned}$$

39

Annexe : paramètres de performance M/M/K



Distribution du temps d'attente dans la file :

- Soit T le temps d'attente d'un client qui arrive dans la file et Q le nombre de clients dans la file à son arrivée. S'il y a un serveur libre, $T=0$. Sinon, il doit attendre que tous les clients de la file soient servis et qu'un serveur se libère.

$$P[T > t] = P[T > t | T > 0] P[T > 0]$$

$$P[T > 0] = E_{2,S}(A)$$

$$P[T > t | T > 0] P[T > 0] = \sum_{j \geq 0} P[T > t | T > 0, Q = j] P[Q = j | T > 0]$$

- La distribution du nombre de clients aux temps d'arrivée est égale à la distribution stationnaire (PASTA) :

$$\begin{aligned}
 P[Q = j | T > 0] &= P[Q = j, T > 0] / P[T > 0] \\
 &= p(j + S) / E_{2,S}(A)
 \end{aligned}$$

38

Glossaire



CCDF : Complementary Cumulative Distribution Function
 CCS : Cent Call Second
 FIFO : First In First Out
 MMPP : Markov Modulated Poisson Process
 RXA : Récepteur d'abonnés

40

Références



- [1] Principes de commutation numérique, Claude Rigault, Éditions Hermès.
- [2] Systèmes et réseaux de télécommunication en régime stochastique, Gérald Doyon, Masson.
- [3] Théorie des files d'attente, Bruno Baynat, Hermès.