

システムアーキテクチャ

M/M/1の待ち行列モデル

M/M/1の意味

M / M / 1

| | └ サービスチャネル数 (サーバ数)

| └ サービス時間分布

└ 到着間隔分布

1つ目のM：到着間隔が Markov 過程 (指数分布)

→ 顧客やパケットが「ポアソン過程」で到着

2つ目のM：サービス時間も 指数分布

最後の1：サーバが 1 台

◆ 前提条件

- 到着はポアソン過程 (平均到着率 = λ)
- サービス時間は指数分布 (平均サービス率 = μ)
- サービスは先着順 (FIFO)
- 系の容量は無限 (待ち人数に制限なし)

◆ 主要なパラメータ

- λ ：到着率 (1 単位時間あたりの平均到着数)
- μ ：サービス率 (1 単位時間あたりに処理できる平均数)
- $\rho = \lambda / \mu$ ：利用率 (サーバの稼働率)
 - $\rho < 1$ のとき、システムは安定 (行列が無限に伸びない)

パラメータについて

1, 平均系内人数 (L)

$$L = \rho / 1 - \rho$$

2, 平均待ち人数 (L_q)

$$L_q = \rho^2 / 1 - \rho$$

3, 平均系内滞在時間 (W)

$$W = 1 / \mu - \lambda$$

4, 平均待ち時間 (W_q)

$$W_q = \rho / \mu - \lambda$$

リトルの公式 ($L = \lambda W$, $L_q = \lambda W_q$) で関係づけられる

覚えにくいので、 $L = \rho / 1 - \rho$ と $W = 1 / \mu - \lambda$ だけ覚えて、それぞれに ρ をかけると、街の値になると覚えるしかない。