On Optimal Policies for Energy-Aware Servers

Vincent J. Maccio

Department of Computing and Software
McMaster University
Hamilton, Ontario, Canada

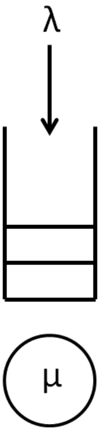
Outline

- Motivation and Introduction
- Preliminary Knowledge
  - $M/M/1$ Queue
  - $M/G/1$ Queue
- Analysis
  - Problem Definition
  - Model
  - Solution
- Applications
- Conclusion

Motivation

1. Data centers and server farms use a lot of energy
2. Data centers and server farms are designed for peak loads
3. Individual servers often idle
4. Idle servers still use a lot of energy

Queueing Theory



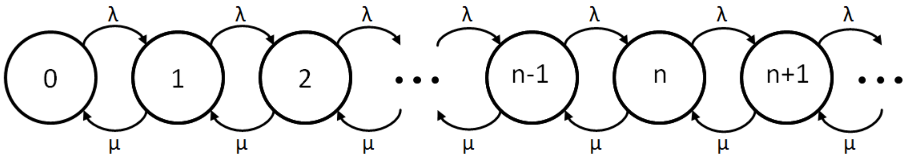
The Exponential Distribution

- Parametrized by some value λ
- pdf: $f(x) = \lambda e^{-\lambda x}$
- CDF: $F(x) = 1 - e^{-\lambda x}$
- Given X is a random variable which is exponentially distributed, it holds that $\forall x, t \geq 0 : P[X > x + t | X > x] = P[X > t]$

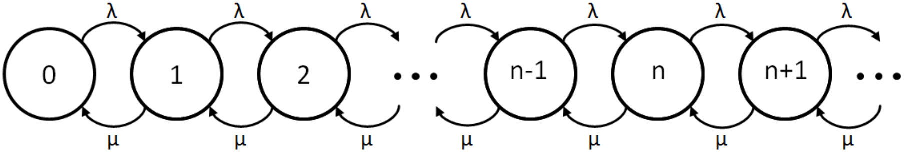
The M/M/1 Queue

- Single server and infinite buffer
- Jobs arrive following a Poisson process with rate λ (time between job arrivals are exponentially distributed with rate λ)
- Job service times are exponentially distributed with rate μ
- Assumed $\mu > \lambda$
- The system utilization is denoted by $\rho = \frac{\lambda}{\mu}$

M/M/1 - Continuous-Time Markov Chain

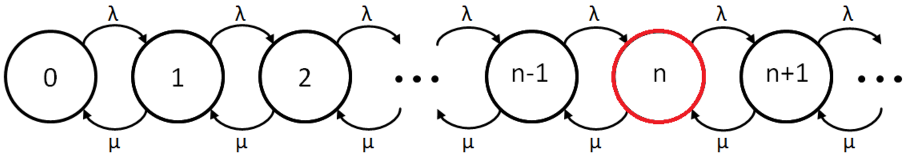


M/M/1 - Analysis



- We wish to arrive at $\mathbb{E}[N]$ and $\mathbb{E}[R]$
- Use steady state analysis

M/M/1 - Analysis



$$n > 0$$

$$\lambda \pi_{n-1} + \mu \pi_{n+1} = (\lambda + \mu) \pi_n$$

$$\pi_1 = \frac{\lambda}{\mu} \pi_0$$

M/M/1 - Analysis

$$n > 0, \qquad \lambda\pi_{n-1} + \mu\pi_{n+1} = (\lambda+\mu)\pi_n, \qquad \pi_1 = \frac{\lambda}{\mu}\pi_0 = \rho\pi_0$$

$$\pi_2 = \frac{\lambda+\mu}{\mu}\pi_1 - \frac{\lambda}{\mu}\pi_0$$

$$\pi_2 = \pi_0\frac{\lambda}{\mu}\left(\frac{\lambda+\mu}{\mu} - 1\right)$$

$$\pi_2 = \rho^2\pi_0$$

$$\pi_n = \rho^n\pi_0$$

M/M/1 - Analysis

$$\sum_{n=0}^{\infty} \pi_n = 1$$

$$\pi_0 \sum_{n=0}^{\infty} \rho^n = 1$$

$$\pi_0 = 1 - \rho$$

$$\pi_n = (1 - \rho)\rho^n$$

M/M/1 - Analysis

With the steady state distribution known, one can arrive at the expected number of jobs in the system, $\mathbb{E}[N]$.

$$\mathbb{E}[N] = \sum_{n=0}^{\infty} n\pi_n \quad \dots \quad \mathbb{E}[N] = \frac{\lambda}{\mu - \lambda}$$

M/M/1 - Analysis

While it is worthwhile knowing the expected number of jobs in the system, it is often more worthwhile to know the expected response time of a job.

$$\mathbb{E}[N] = \lambda \mathbb{E}[R]$$

Therefore,

$$\mathbb{E}[R] = \frac{1}{\mu - \lambda}$$

The M/G/1 Queue

- Single server and infinite buffer
- Jobs arrive following a Poisson process with rate λ (time between job arrivals are exponentially distributed with rate λ)
- Job service times are generally distributed with rate μ
- Assumed $\mu > \lambda$
- Denote the system utilization $\rho = \frac{\lambda}{\mu}$

M/G/1 Analysis

- The system can no longer be viewed as a CTMC
- However, an embedded Markov chain exists
- Let N_n be a random variable denoting the number of jobs in the system when the n^{th} job leaves
- Let A_n be a random variable denoting the number of jobs which arrived during the n^{th} service time

M/G/1 Analysis

$$N_{n+1} = \begin{cases} N_n - 1 + A_{n+1} & N_n > 0 \\ A_{n+1} & N_n = 0 \end{cases}$$

$$\mathcal{U}(x) = \begin{cases} 1 & x > 0 \\ 0 & x = 0 \end{cases}$$

$$N_{n+1} = N_n - \mathcal{U}(N_n) + A$$

M/G/1 Analysis

$$N_{n+1}^2 = N_n^2 + \mathcal{U}^2(N_n) + A^2 - 2N_n\mathcal{U}(N_n) - 2A\mathcal{U}(N_n) + 2N_nA$$

$$\mathcal{U}^2(N_n) = \mathcal{U}(N_n), \qquad N_n\mathcal{U}(N_n) = N_n$$

$$0 = \mathbb{E}[\mathcal{U}(N_n)] + \mathbb{E}[A^2] - 2\mathbb{E}[N_n] - 2\mathbb{E}[A]\mathbb{E}[\mathcal{U}(N_n)] + 2\mathbb{E}[N_n]\mathbb{E}[A]$$

$$\mathbb{E}[\mathcal{U}(N_n)] = \mathbb{E}[A] = \rho, \qquad \mathbb{E}[A^2] = \rho + \lambda^2\sigma_S^2$$

M/G/1 Analysis

After some algebra, the $M/G/1$ queue is solved for.

$$\mathbb{E}[N] = \rho + \frac{\rho^2 + \lambda^2 \sigma_S^2}{2(1-\rho)}$$

$$\mathbb{E}[N] = \frac{1}{\mu} + \frac{\rho^2 + \lambda^2 \sigma_S^2}{2\lambda(1-\rho)}$$

Goal

- Derive optimal policies, when energy used by the system is a concern
- Two decisions, when to turn it off, when to turn it on

Previous Work

- Gandhi et al. [1], [3] determined a set of optimal policies for single server system under the cost function $\mathbb{E}[R]\mathbb{E}[E]$. Solved steady state distribution for a server which instantly turns off when it idles, and begins to turn back on once a job arrives.
- Various authors [4], [5], [6], [7] use a cost function of the form $\mathbb{E}[R] + \beta_1\mathbb{E}[E] + \beta_2\mathbb{E}[Sw]$, with great analysis. However, optimal policies remained unknown (even for single server).
- Various authors [7], [8] examine multi-server systems under certain policies. However, optimal policies remain unknown.

Our Contribution

Give one the tools to derive the optimal policy for a single on/off server system under any cost function of the form:

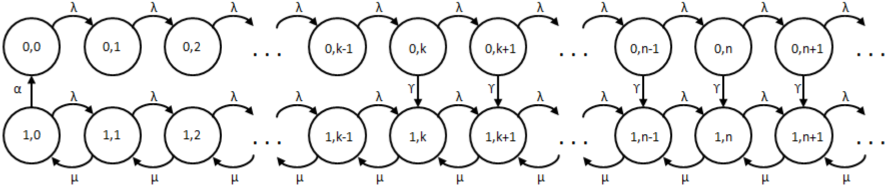
$$f(\beta, w) = \sum_{i=1}^{M} \beta_i \mathbb{E}[R]^{w_{R,i}} \mathbb{E}[E]^{w_{E,i}} \mathbb{E}[Sw]^{w_{Sw,i}},$$

where all $\beta, w \geq 0$.

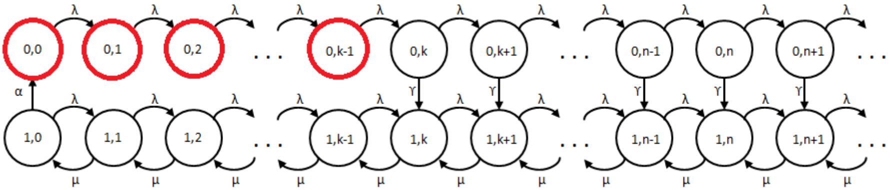
Model

- Single server, infinite queue
- Arrivals follow Poisson process, rate λ
- Service times exponentially distributed, rate μ
- Server can be in one of four (energy) states - *OFF*, *BUSY*, *IDLE*, *SETUP*
- Corresponding energy values, $E_{Off} = 0$, E_{Setup}, E_{Busy} and E_{Idle}
- Turn on time, exponentially distributed, rate γ
- Time to wait before moving to low state, exponentially distributed, rate α (idling times)
- Time needed to turn off equals zero
- Server in low state waits for k arrivals before beginning to turn on

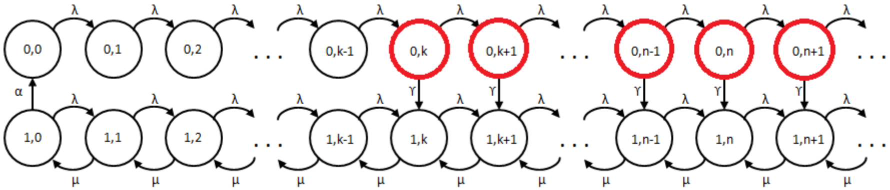
Markov Chain



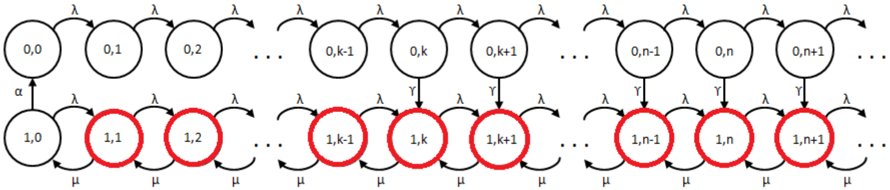
Markov Chain - Off



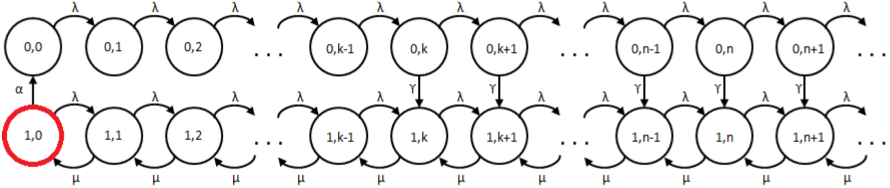
Markov Chain - Setup



Markov Chain - Busy



Markov Chain - Idle



Properties of Optimal Policy

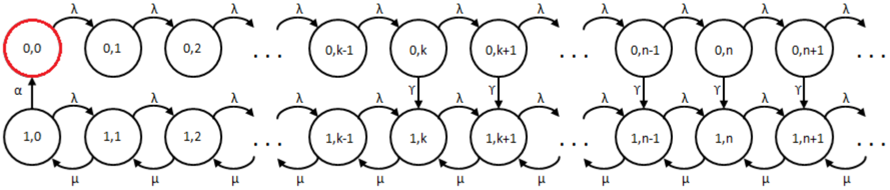
The following properties carry over from Gandhi et al. [1]:

- The decision to start transitioning between lower and higher energy states is made at the moment a job arrives to the system.
- When the server idles, it is always optimal to immediately turn the server off, or to always keep the server on.
- If there are jobs in the system and the server is in its higher energy state, the server will never move to its lower energy state.

Markov Chain - Solution

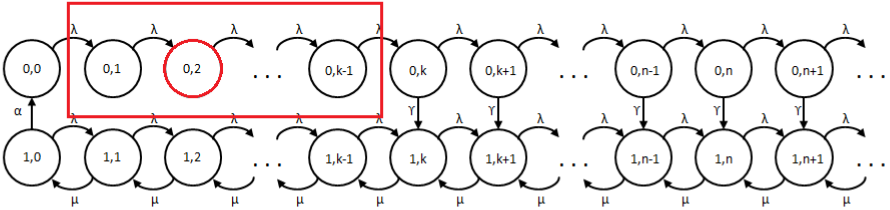
- Look for patterns and repeating portions
- Pick intelligent balance and boundary equations
- Do the math

Markov Chain - Boundary Equation



$$\alpha\pi_{1,0} = \lambda\pi_{0,0} \quad \Rightarrow \quad \pi_{1,0} = \frac{\lambda}{\alpha}\pi_{0,0}$$

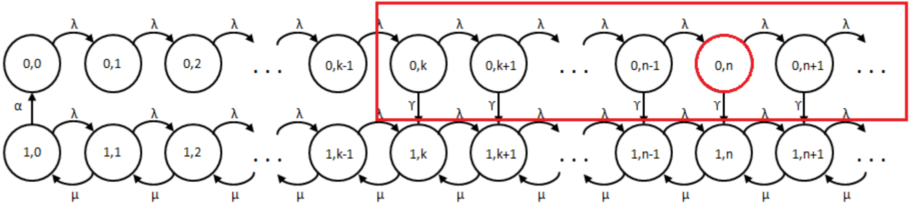
Markov Chain - Balance Equations



$$(0 < n < k)$$

$$\lambda \pi_{0,n-1} = \lambda \pi_{0,n} \quad \Rightarrow \quad \pi_{0,n} = \pi_{0,0}$$

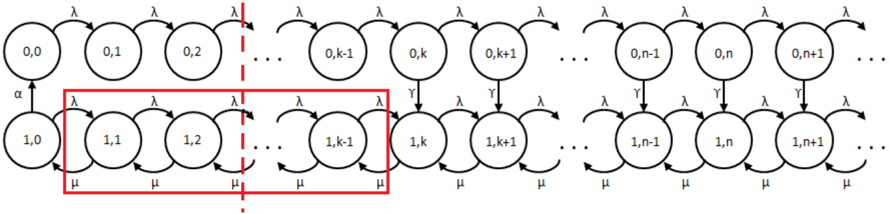
Markov Chain - Balance Equations



$$(k \leq n)$$

$$\lambda \pi_{0,n-1} = (\lambda + \gamma)\pi_{0,n} \quad \Rightarrow \quad \pi_{0,n} = \pi_{0,0}\left(\frac{\lambda}{\lambda+\gamma}\right)^{n-(k-1)}$$

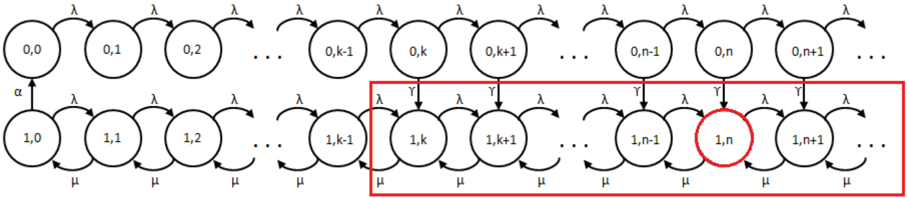
Markov Chain - Balance Equations



$$(0 < n < k)$$

$$\lambda(\pi_{0,n-1} + \pi_{1,n-1}) = \mu\pi_{1,n} \quad \Rightarrow \quad \pi_{1,n} = \pi_{0,0}\left(\frac{\lambda}{\alpha}\rho^n + \frac{\lambda}{\mu - \lambda}(1 - \rho^n)\right)$$

Markov Chain - Balance Equations



$$(k \leq n)$$

$$\lambda\pi_{1,n-1} + \mu\pi_{1,n+1} + \gamma\pi_{0,n} = (\mu + \lambda)\pi_{1,n} \quad \Rightarrow \quad \pi_{1,n} = ?$$

Markov Chain - Balance Equations

$$\pi_{1,n} = Ax^{(n-(k-1))} + B\left(\frac{\lambda}{\lambda+\gamma}\right)^{n-(k-1)}$$

where,

$$(\lambda+\mu)x = \lambda + \mu x^2 \quad \Rightarrow \quad x = 1, \rho$$

$$B = \pi_{0,0}\frac{\lambda+\gamma}{\mu-\lambda-\gamma}$$

$$A = \pi_{0,0}\left[\left(\frac{\lambda}{\alpha} - \frac{\lambda}{\mu-\lambda}\right)\rho^{k-1} - \frac{\mu\gamma}{(\mu-\lambda)(\mu-\lambda-\gamma)}\right]$$

Markov Chain - Balance Equations

$$\pi_{1,n} = \pi_{0,0}\left[\left(\frac{\lambda}{\alpha} - \frac{\lambda}{\mu-\lambda}\right)\rho^{n} + \frac{1}{\mu-\lambda-\gamma}\left((\lambda+\gamma)\left(\frac{\lambda}{\lambda+\gamma}\right)^{n-(k-1)} - \frac{\gamma}{1-\rho}\rho^{n-(k-1)}\right)\right]$$

Steady-State Distribution

$$
\begin{aligned}
\pi_{0,n} &= \pi_{0,0}, \qquad 0 \le n < k \\
\pi_{0,n} &= \pi_{0,0} \left(\frac{\lambda}{\lambda+\gamma} \right)^{n-(k-1)}, \qquad k \le n \\
\pi_{1,n} &= \pi_{0,0} \left(\frac{\lambda}{\alpha} \rho^n + \frac{\lambda}{\mu-\lambda} (1-\rho^n) \right), \qquad 0 \le n < k \\
\pi_{1,n} &= \pi_{0,0} \left[\left(\frac{\lambda}{\alpha} - \frac{\lambda}{\mu-\lambda} \right) \rho^n + \frac{1}{\mu-\lambda-\gamma} \left((\lambda+\gamma) \left(\frac{\lambda}{\lambda+\gamma} \right)^{n-(k-1)} - \frac{\gamma}{1-\rho} \rho^{n-(k-1)} \right) \right], \quad k \le n \\
\pi_{0,0} &= (1-\rho) \frac{\alpha\gamma}{k\alpha\gamma + \alpha\lambda + \lambda\gamma}
\end{aligned}
$$

The Three Metrics

$$\mathbb{E}[Sw] = (1-\rho)\frac{\alpha\lambda\gamma}{k\alpha\gamma+\alpha\lambda+\lambda\gamma}$$

$$\mathbb{E}[E] = \mathbb{E}[E_{M/M/1}] + \frac{(1-\rho)\alpha}{k\alpha\gamma+\alpha\lambda+\lambda\gamma}(\lambda E_{Setup} - (\lambda+k\gamma)E_{Idle})$$

$$\mathbb{E}[R] = \mathbb{E}[R_{M/M/1}] + \frac{1}{\gamma}\frac{\alpha(\lambda+k\gamma)}{k\alpha\gamma+\alpha\lambda+\lambda\gamma} + \frac{1}{2\lambda}\frac{k\alpha\gamma(k-1)}{k\alpha\gamma+\alpha\lambda+\lambda\gamma}$$

	Optimal Values of	
Metric	α	k
$\mathbb{E}[R]$	0	1
$\mathbb{E}[E]$	0 or $\to\infty$	$\to\infty$
$\mathbb{E}[Sw]$	0	$\to\infty$

Weighted Sum Cost Function

$$\mathbb{E}[R] + \beta_1 \mathbb{E}[E] + \beta_2 \mathbb{E}[Sw]$$

Let $r_{idle} = E_{Idle}/E_{Busy}$ and $r_{setup} = E_{Setup}/E_{Busy}$.

- Choose $\alpha = \infty$ if

$$\beta_1(1-\rho)\lambda\gamma(k\gamma+\lambda) \leq \lambda(\lambda+k\gamma)+\gamma k(k-1)+\beta_1(1-\rho)\lambda^2\gamma r_{setup}+\beta_2(1-\rho)\lambda^2\gamma^2$$

  Otherwise, choose $\alpha = 0$.

- Solve

$$0 = \frac{\gamma^2}{2\lambda}k^2 + \gamma k - (\lambda + (1-\rho)\lambda\gamma(\beta_1 r_{setup} + \beta_2\gamma))$$

  If a valid solution, check $\lceil k^* \rceil$ and $\lfloor k^* \rfloor$.

The Work-Cycle

- Let OFF_0 denote the state where there are no jobs in the system, and the system is in the energy state OFF
- Let $P_{Off,0}$ denote the steady state probability of being in state OFF_0
- A Work-Cycle is the system starting in state OFF_0, leaving that state, and then returning to it
- The system can be viewed as a series of Work-Cycles

Work-Cycle Properties

- During each Work-Cycle, the system is in energy state *SETUP* for some amount of time, expected to be $\frac{1}{\gamma}$
- The system is in the energy state *OFF* while k jobs arrive, expected to take $\frac{k}{\lambda}$ amount of time
- The system is in the energy state *IDLE* for some amount of time, expected to be $\frac{1}{\alpha}$
- The amount of time spent in the energy state *BUSY* during a single Work-Cycle is not clear...
- The rate at which Work-Cycles occur is equal to $\lambda P_{Off,0}$
- The only assumption which has been made on the underlying distributions, is that jobs arrive according to a Poisson process

Work-Cycle Analysis

- Let P_{Off}, P_{Setup}, P_{Busy}, and P_{Idle} denote the steady state probabilities of the system being in the energy states *OFF*, *SETUP*, *BUSY*, and *IDLE* respectively
- It's observed that

$$
\begin{aligned}
P_{Off} &= \frac{k\lambda}{\lambda} P_{Off,0} = kP_{Off,0} \\
P_{Setup} &= \frac{\lambda}{\gamma} P_{Off,0} \\
P_{Idle} &= \frac{\lambda}{\alpha} P_{Off,0}
\end{aligned}
$$

Work-Cycle Analysis

- From the Work-Cycle viewpoint finding P_{Busy} could be challenging, however, it is known that $P_{Busy} = \rho$
- It is also observed that

$$P_{Off} + P_{Setup} + P_{Busy} + P_{Idle} = 1$$
$$kP_{Off,0} + \frac{\lambda}{\gamma} P_{Off,0} + \frac{\lambda}{\alpha} P_{Off,0} = 1 - \rho$$
$$(1-\rho)\frac{\alpha\gamma}{k\alpha\gamma + \alpha\lambda + \lambda\gamma} = P_{Off,0}$$

$$\Rightarrow \quad P_{Off,0} = \pi_{0,0}$$

Work-Cycle Analysis

With $P_{Off,0}$ solved, the steady state probabilities for the energy states could be determined.

$$\mathbb{E}[E] = \mathbb{E}[E_{M/G/1}] + \frac{(1-\rho)\alpha}{k\alpha\gamma + \alpha\lambda + \lambda\gamma}(\lambda E_{Setup} - (\lambda + k\gamma)E_{Idle})$$

The expected switching rate of the server, is equal to the rate of the Work-Cycles, $\lambda P_{Off,0}$.

$$\mathbb{E}[Sw] = (1-\rho)\frac{\alpha\lambda\gamma}{k\alpha\gamma + \alpha\lambda + \lambda\gamma}$$

Generalizing Distributional Assumptions

- Allowing all distributions to be arbitrary, other than the arrival process being Poisson, $\mathbb{E}[E]$ and $\mathbb{E}[Sw]$ are the same as for the setting with all distributions exponential.
- $\mathbb{E}[E]$ and $\mathbb{E}[Sw]$ are dependent only on the means of the underlying distributions.
- If $r_{idle} < \frac{\lambda}{k\gamma+\lambda} r_{setup}$ it is always optimal to leave the server on.

Generalizing Distributional Assumptions Analysis

- If the arrival stream remains Poisson, and the idling time of the system remains exponentially distributed, can one derive the mean response time using M/G/1-type arguments (embedded chain)?
- If a closed form expression can be derived, the optimal policy can be determined under general assumptions

Generalizing Distributional Assumptions Analysis

Using a similar recursive definition from the classical $M/G/1$ analysis.

$$N_{n+1} = \begin{cases} N_n + A_{n+1} - 1 & N_n \geq 1 \\ A_{n+1} & N_n = 0 \end{cases}$$

$$A_{n+1} = \begin{cases} A_{S,n+1} & N_n \geq 1 \\ A_{S,n+1} + X_{Off,n}(k - 1 + A_{\Gamma,n}) & N_n = 0 \end{cases}$$

Generalizing Distributional Assumptions Analysis

$$
\begin{aligned}
\mathbb{E}[N^2] = \mathbb{E}[N^2] &- 2\mathbb{E}[N\mathcal{U}(N)] + 2\mathbb{E}[NA_S] \\
&+ \mathbb{E}[N(1-\mathcal{U}(N))X_{Off}(k-1+A_\Gamma)] \\
&+ \mathbb{E}[\mathcal{U}^2(N)] - 2\mathbb{E}[\mathcal{U}(N)A_S] \\
&- 2\mathbb{E}[\mathcal{U}(N)(1-\mathcal{U}(N))X_{Off}(k-1+A_\Gamma)] \\
&+ \mathbb{E}[A_S^2] + 2\mathbb{E}[A_S(1-\mathcal{U}(N))X_{Off}(k-1+A_\Gamma)] \\
&+ \mathbb{E}[(1-\mathcal{U}(N))^2 X_{Off}^2 (k-1+A_\Gamma)^2]
\end{aligned}
$$

Generalizing Distributional Assumptions Analysis

After some algebra, a closed form equation can be derived,

$$\mathbb{E}[N] = \mathbb{E}[N_{M/G/1}] + \frac{\alpha}{\alpha+\lambda}\left[\rho\frac{(k-1)\gamma+\lambda}{\gamma} + \frac{1}{2}\Gamma\right] + \alpha\frac{(k-1)\gamma+\lambda}{k\alpha\gamma+\alpha\lambda+\lambda\gamma}\left[\frac{1}{2} - \rho - \rho\frac{\alpha}{\alpha+\lambda}\left(\frac{(k-1)\gamma+\lambda}{\gamma}\right) - \frac{1}{2}\frac{\alpha}{\alpha+\lambda}\Gamma\right]$$

where,

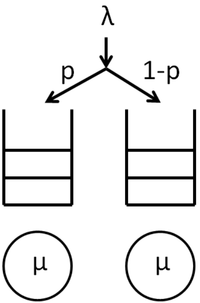
$$\Gamma = (k-1)^2 + (2k-1)\frac{\lambda}{\gamma} + \lambda^2\sigma_\Gamma^2$$

Constrained Optimization

Minimize the expected energy subject to $\mathbb{E}[R] \leq T$, where $\frac{1}{\mu-\lambda} \leq T \leq \frac{1}{\mu-\lambda} + \frac{1}{\gamma}$. Fix $k = 1$.

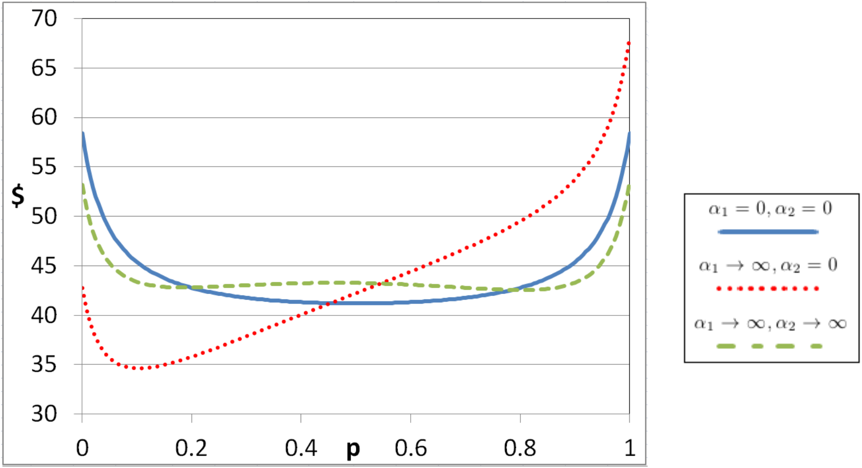
$$\alpha = \frac{\lambda\gamma^2}{\lambda+\gamma} \frac{T - \mathbb{E}[R_{M/M/1}]}{1 - \gamma(T - \mathbb{E}[R_{M/M/1}])}$$

Random Routing



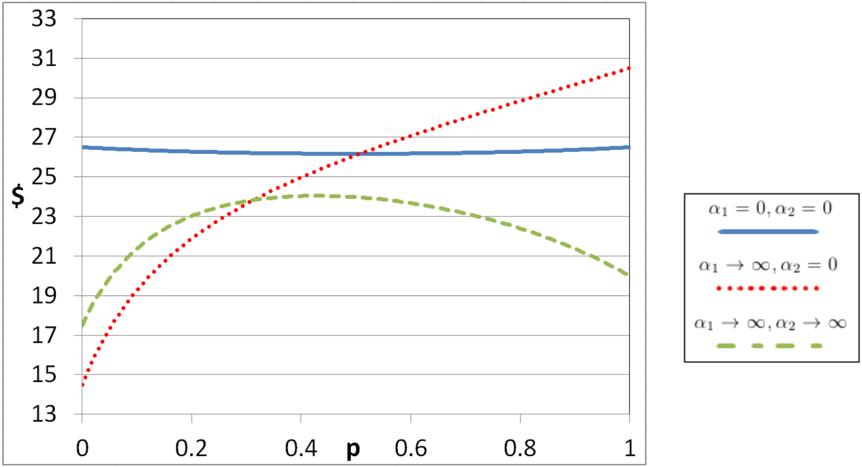
- Five parameters to choose: α_1, α_2, k_1, k_2, and p.
- Minimize $\mathbb{E}[N] + \beta\mathbb{E}[E]$

With $\mu = 2$, $\lambda = 1.9$, $\gamma = 0.2$, $\beta = 22$



Here $k_1 = 5$ for the optimal configuration and $p = 0.105$.

With $\mu = 2$, $\lambda = 1$, $\gamma = 0.2$, $\beta = 15$



Conclusion

- Complete analysis of single server systems.
- In random routing scenario, heterogeneous servers may be desirable.

Future Work

- Add more options and configurations to the single server model (sleep states, speed scaling etc.)
- Look at the multi-server problem (what properties does the optimal policy have in this setting?)

[1] A. Gandhi, V. Gupta, M. Harchol-Balter, and M. Kozuch. "Optimality Analysis of Energy-Performance Trade-off for Server Farm Management." *Performance Evaluation*, vol. 67, 2010, pp. 1155-1171.

[2] A. Gandhi and M. Harchol-Balter. "M/M/k with Exponential Setup." *CMU Technical Report CMU-CS-09-166*, 2010.

[3] A. Gandhi, M. Harchol-Balter, and I. Adan. "Server Farms with Setup Costs." *Performance Evaluation*, vol. 67, no. 11, 2010, pp. 1123-1138.

[4] A. Penttinen, E. Hyytia, and S. Aalto. "Energy-Aware Dispatching in Parallel Queues with On-Off Energy Consumption." *IEEE International Performance Computing and Communications Conference*, 2011.

[5] Y. Chen, A. Das, W. Qin, A. Sivasubramaniam, Q. Wand, and N. Gautam. "Managing Server Energy and Operational Costs in Hosting Centers." *ACM SIGMETRICS*, 2005, pp. 303-314.

[6] A. Wierman, L. L. H. Andrew, and M. Lin. "Speed Scaling: An Algorithmic Perspective." *Chapter in Handbook on Energy-Aware and Green Computing*, CRC Press, 2012, pp. 385-406.

[7] J. Slegers, N Thomas, and I. Mitrani. "Dynamic Server Allocation for Power and Performance." *Performance Evaluation: Metrics, Models and Benchmarks Lecture Notes in Computer Science*, vol. 5119, 2008, pp. 247-261.

[8] A. Gandhi, S. Doroudi, M. Harchol-Balter and A. Scheller-Wolf. "Exact Analysis of the M/M/k/setup Class of Markov Chains via Recursive Renewal Reward." *ACM SIGMETRICS*, 2013.

Questions?

Thank you.