Stability of Queueing Networks with Infinite Supplies

Yoni Nazarathy,
The University of Queensland,

Based on some joint work with

Erjen Lefeber, Eindhoven University of Technology,
Gideon Weiss, The University of Haifa and The University of Southern California,
Hanqin Zhang, National University of Singapore.

QANZIAM,
October 28, 2012.

The Study of Queueing Networks in Applied Probability and Operations Research

Congestion, delay and resource scarcity occurs
in a variety of application areas:

- Customer service systems
- Complex manufacturing lines
- Telecommunication networks and computing systems
- Transportation networks

The Study of Queueing Networks in Applied Probability and Operations Research

Congestion, delay and resource scarcity occurs in a variety of application areas:

- Customer service systems
- Complex manufacturing lines
- Telecommunication networks and computing systems
- Transportation networks

Stochastic queueing network models capture the essential behaviour, allowing for quantitative performance evaluation, optimization and control

A Single Queue

Items arrive at random times to a server, queue up, each requiring service for a random duration, then depart

A Single Queue

Items arrive at random times to a server, queue up, each requiring service for a random duration, then depart

Construction of the Queue Length Process: $Q(t)$

$A(t) \equiv$ counting process generated by a sequence of random **inter-arrival times** each with mean λ^{-1}

$S(t) \equiv$ counting process generated by a sequence of random **service times** each with mean μ^{-1}

A Single Queue

Items arrive at random times to a server, queue up, each requiring service for a random duration, then depart

Construction of the Queue Length Process: $Q(t)$

$A(t) \equiv$ counting process generated by a sequence of random **inter-arrival times** each with mean λ^{-1}

$S(t) \equiv$ counting process generated by a sequence of random **service times** each with mean μ^{-1}

$$\begin{aligned} Q(t) &= Q(0) + A(t) - S\big(T(t)\big) \\ T(t) &= \int_0^t \mathbf{1}_{\{Q(s)>0\}} ds \end{aligned}$$

A Single Queue

Items arrive at random times to a server, queue up, each requiring service for a random duration, then depart

Construction of the Queue Length Process: $Q(t)$

$A(t) \equiv$ counting process generated by a sequence of random **inter-arrival times** each with mean λ^{-1}

$S(t) \equiv$ counting process generated by a sequence of random **service times** each with mean μ^{-1}

$$
\begin{aligned}
Q(t) &= Q(0) + A(t) - S\big(T(t)\big) \\
T(t) &= \int_0^t \mathbf{1}_{\{Q(s)>0\}} ds
\end{aligned}
$$

The Load $\rho = \frac{\lambda}{\mu}$

A Single Queue

Items arrive at random times to a server, queue up, each requiring service for a random duration, then depart

Construction of the Queue Length Process: $Q(t)$

$A(t) \equiv$ counting process generated by a sequence of random **inter-arrival times** each with mean λ^{-1}

$S(t) \equiv$ counting process generated by a sequence of random **service times** each with mean μ^{-1}

$$
\begin{aligned}
Q(t) &= Q(0) + A(t) - S\big(T(t)\big) \\
T(t) &= \int_0^t \mathbf{1}_{\{Q(s)>0\}} ds
\end{aligned}
$$

The Load $\rho = \frac{\lambda}{\mu}$

- $\rho < 1$ queue is **stable**

A Single Queue

Items arrive at random times to a server, queue up, each requiring service for a random duration, then depart

Construction of the Queue Length Process: $Q(t)$

$A(t) \equiv$ counting process generated by a sequence of random **inter-arrival times** each with mean λ^{-1}

$S(t) \equiv$ counting process generated by a sequence of random **service times** each with mean μ^{-1}

$$
\begin{aligned}
Q(t) &= Q(0) + A(t) - S\big(T(t)\big) \\
T(t) &= \int_0^t \mathbf{1}_{\{Q(s)>0\}} ds
\end{aligned}
$$

The Load $\rho = \frac{\lambda}{\mu}$

- $\rho < 1$ queue is **stable**
- $\rho > 1$ queue is **unstable**

A Single Queue

Items arrive at random times to a server, queue up, each requiring service for a random duration, then depart

Construction of the Queue Length Process: $Q(t)$

$A(t) \equiv$ counting process generated by a sequence of random **inter-arrival times** each with mean λ^{-1}

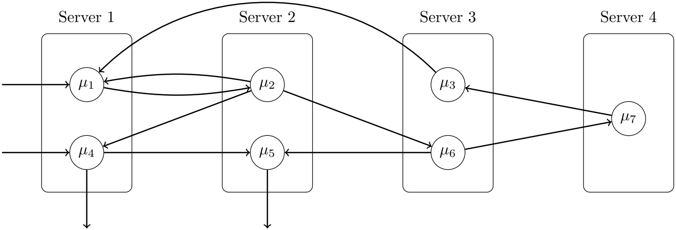
$S(t) \equiv$ counting process generated by a sequence of random **service times** each with mean μ^{-1}

$$\begin{aligned} Q(t) &= Q(0) + A(t) - S\big(T(t)\big) \\ T(t) &= \int_0^t \mathbf{1}_{\{Q(s)>0\}} ds \end{aligned}$$

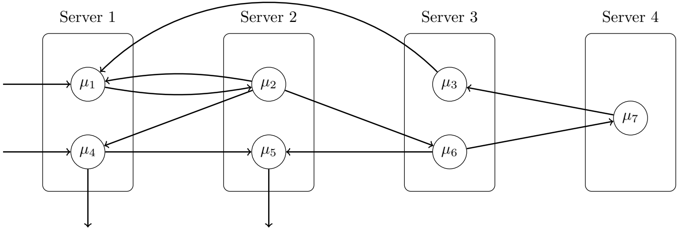
The Load $\rho = \frac{\lambda}{\mu}$

- $\rho < 1$ queue is **stable**
- $\rho > 1$ queue is **unstable**
- $\rho = 1$ queue is **critically unstable**

Multi-Class Queueing Networks



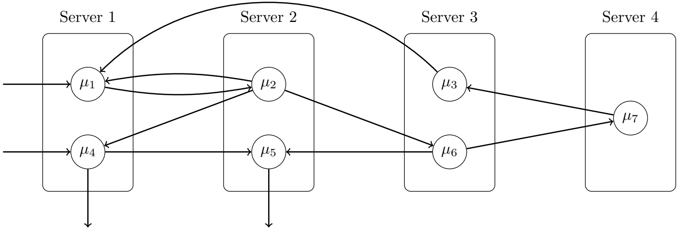
Multi-Class Queueing Networks



Feedback Control Policies

Policy: What operation should be served by each of the servers at every time instant based on the current state

Multi-Class Queueing Networks

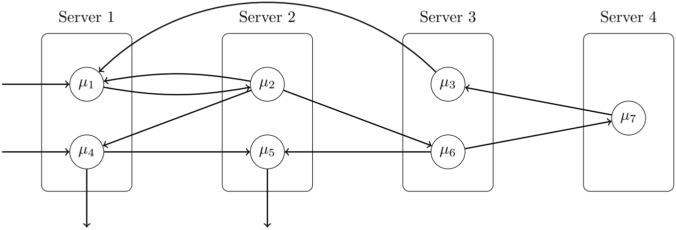


Feedback Control Policies

Policy: What operation should be served by each of the servers at every time instant based on the current state

- Explicit performance analysis for a given control policy is typically intractable

Multi-Class Queueing Networks

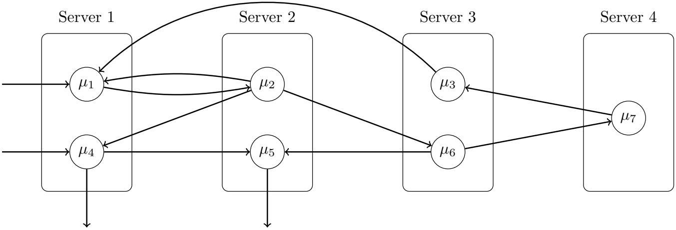


Feedback Control Policies

Policy: What operation should be served by each of the servers at every time instant based on the current state

- Explicit performance analysis for a given control policy is typically intractable
- Optimal control is typically out of the question

Multi-Class Queueing Networks



Feedback Control Policies

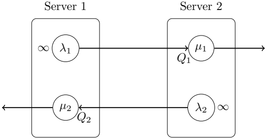
Policy: What operation should be served by each of the servers at every time instant based on the current state

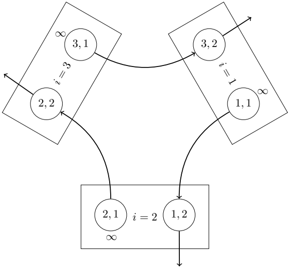
- Explicit performance analysis for a given control policy is typically intractable
- Optimal control is typically out of the question

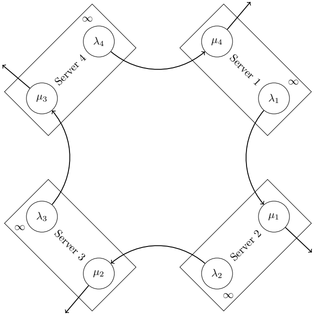
A realistic research goal:
understanding stability for given policies

Queueing Networks with Infinite Supplies

Push-Pull Rings with Pull-Priority Policy







Stochastic Model and Fluid Model

Stochastic Model $(Q(t), T(t))$

$$\begin{aligned} Q_i(t) &= Q_i(0) + S_{i,1}\big(T_{i,1}(t)\big) - S_{i,2}\big(T_{i,2}(t)\big) \\ t &= T_{i,1}(t) + T_{i-1,2}(t) \\ 0 &= \int_0^t Q_i(s)\,dT_{i+1,1}(s) \end{aligned}$$

Stochastic Model and Fluid Model

Stochastic Model $\big(Q(t), T(t)\big)$

$$
\begin{aligned}
Q_i(t) &= Q_i(0) + S_{i,1}\big(T_{i,1}(t)\big) - S_{i,2}\big(T_{i,2}(t)\big) \\
t &= T_{i,1}(t) + T_{i-1,2}(t) \\
0 &= \int_0^t Q_i(s) dT_{i+1,1}(s)
\end{aligned}
$$

Associated Fluid Model $\big(\bar{Q}(t), \bar{T}(t)\big)$

$$
\begin{aligned}
\bar{Q}_i(t) &= \bar{Q}_i(0) + \lambda_i \bar{T}_{i,1}(t) - \mu_i \bar{T}_{i,2}(t) \\
t &= \bar{T}_{i,1}(t) + \bar{T}_{i-1,2}(t) \\
0 &= \int_0^t \bar{Q}_i(s) d\bar{T}_{i+1,1}(s)
\end{aligned}
$$

Fluid Stability Framework

Thm: (Dai '95), adapted to infinite supplies

Assume minor technical assumptions on the processing time distributions. If there exists a τ such that for all solutions of the fluid model and all $t \geq \tau$, $\sum \bar{Q}_i(t) = 0$ then the (stochastic) network is stable.

- All solutions of the fluid model are Lipschitz, thus have derivatives a.e.
- **Regular time points**: Time points at which derivatives exists

Lemma

If we have a Lyapounov function: $V : \mathbb{R}^M \rightarrow \mathbb{R}$ such that for all regular time points of all solutions of the fluid model, $\frac{d}{dt} V\left(\bar{Q}(t)\right) < -\epsilon$ for some $\epsilon > 0$, then the fluid model is stable.

Stability Result: M odd, $\gamma_i := \frac{\lambda_i}{\mu_i} > 1$

Theorem (Erjen Lefeber, Gideon Weiss, Y.N.)

The push-pull ring with M **odd**, $\gamma_i > 1$ for all i, operating under a pull-priority policy is stable if $\Delta < 0$, where

$$\Delta = \sum_{i=1}^{M} c_i \Big(\frac{M-1}{2} (\gamma_i - 1) - 1 \Big),$$

with,

$$c_i = ((\cdots(((\gamma_{i-1} - 1)\gamma_{i-2} + 1)\gamma_{i-3} - 1)\gamma_{i-4} \cdots\cdots\cdots)\gamma_{i+2} - 1)\gamma_{i+1} + 1.$$

Note: If $\gamma_i = \gamma$ for all i then the stability condition reduces to:

$$\gamma < 1 + \frac{1}{\frac{M-1}{2}}$$

Proof Overview

1. Use $V(x) = \sum_{i=1}^{M} c_i\, x_i$ as Lyapounov function for the fluid model with coefficients, c_i, designed based on the intuition that "typical" fluid trajectories eventually cycle on states of the form (e.g. $M = 5$):

$$(+,0,+,0,+),\ \ (+,+,0,+,0),\ \ (0,+,+,0,+),\ \ (+,0,+,+,0),\ \ (0,+,0,+,+).$$

The c_i are such that $\dot{V}(t)$ is constant during such cycles:

$$\begin{bmatrix} -1 & 0 & \gamma_3 - 1 & 0 & \gamma_5 - 1 \\ \gamma_1 - 1 & -1 & 0 & \gamma_4 - 1 & 0 \\ 0 & \gamma_2 - 1 & -1 & 0 & \gamma_5 - 1 \\ \gamma_1 - 1 & 0 & \gamma_3 - 1 & -1 & 0 \\ 0 & \gamma_2 - 1 & 0 & \gamma_4 - 1 & -1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \end{bmatrix} = \begin{bmatrix} \Delta \\ \Delta \\ \Delta \\ \Delta \\ \Delta \end{bmatrix}$$

2. Characterize the regular time points and show that on these time points the Lyapounov function has negative drift
3. Now apply Jim Dai's stability framework

The End