OPTIMAL RESOURCE ALLOCATION IN DATA CENTER NETWORKS: DRIFT METHOD AND TRANSFORM TECHNIQUES

Daniela Hurtado-Lange, Siva Theja Maguluri

Industrial and Systems Engineering

Georgia Institute of Technology

September 2019

GEORGIA INSTITUTE OF TECHNOLOGY

Overall

#4 Most Innovative Schools

#5 Top Public Universities

#5 Co-op and Internship Programs

#8 Undergraduate Research / Creative Projects

#10 Best Colleges for Veterans

#12 Senior Capstone Projects

#20 Learning Communities

#29 National Universities

#69 Global Universities

As of September 2019

Undergraduate

#4 Best Undergraduate Engineering Program

- #1 Industrial / Systems Manufacturing Engineering
- #2 Aerospace / Aeronautical / Astronautical Engineering
- #2 Civil Engineering
- #2 Materials Science & Engineering
- #3 Chemical Engineering
- #3 Mechanical Engineering
- #4 Biomedical Engineering
- #4 Computer Engineering
- #4 Electrical / Electronic / Communications Engineering
- #4 Environmental / Environmental Health Engineering

#22 Best Undergraduate Business Program

- #3 Management Information Systems
- #4 Quantitative Analysis
- #7 Production / Operations Management
- #8 Supply Chain Management / Logistics

As of September 2019

Graduate

#7 Best Graduate Engineering School

- #1 Industrial Engineering
- #2 Biomedical Engineering
- #4 Aerospace Engineering
- #4 Civil Engineering
- #5 Computer Engineering
- #4 Environmental Engineering
- #5 Mechanical Engineering
- #6 Electrical Engineering
- #7 Chemical Engineering
- #7 Materials Science & Engineering
- #9 Nuclear Engineering

#8 Best Computer Science Graduate Program

- #7 Artificial Intelligence
- #9 Theory
- #10 Systems

#26 Best Graduate Mathematics School

- #2 Discrete Mathematics and Combinatorics

#29 Best Graduate Business School

- #7 Production / Operations
- #8 Information Systems
- #11 Supply Chain / Logistics

Graduate Sciences Rankings

- #20 Chemistry
- #28 Physics
- #26 Mathematics
- #38 Earth Sciences

As of March 2019

INDUSTRIAL AND SYSTEMS ENGINEERING (ISYE)

ISyE Rankings

#1 Undergraduate Program

- The 25th consecutive No. 1 ranking for ISyE. (USNWR's America's Best Colleges)

#1 Graduate Program

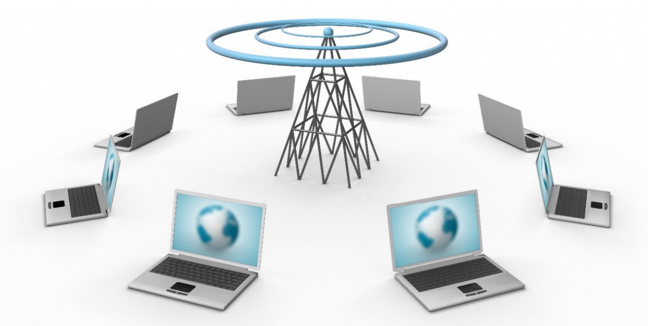
- The 29th consecutive No. 1 ranking for ISyE in the industrial/manufacturing/systems specialty. (USNWR's Best Graduate Programs)

#8 Statistics & Operational Research

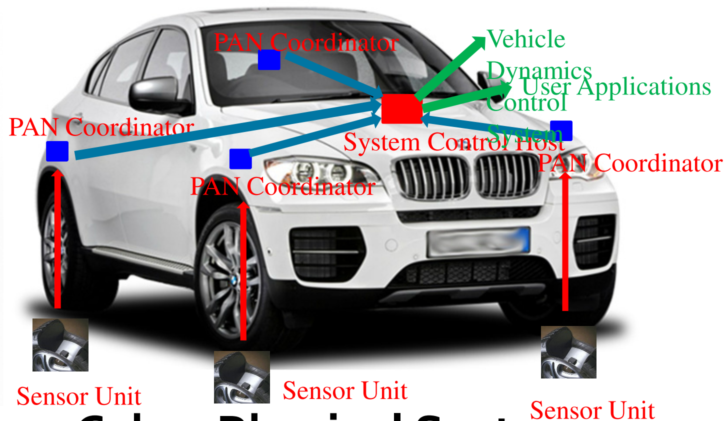
- (2019 QS World University Rankings)

MBA
Master
PhD

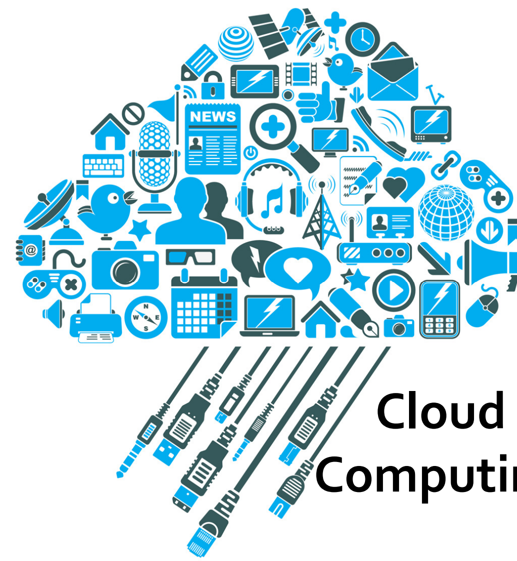
STOCHASTIC PROCESSING NETWORKS



Wireless Networks

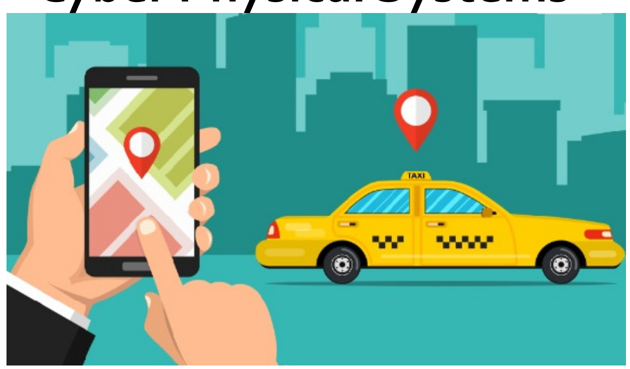


Cyber Physical Systems



Cloud Computing

Data Centers

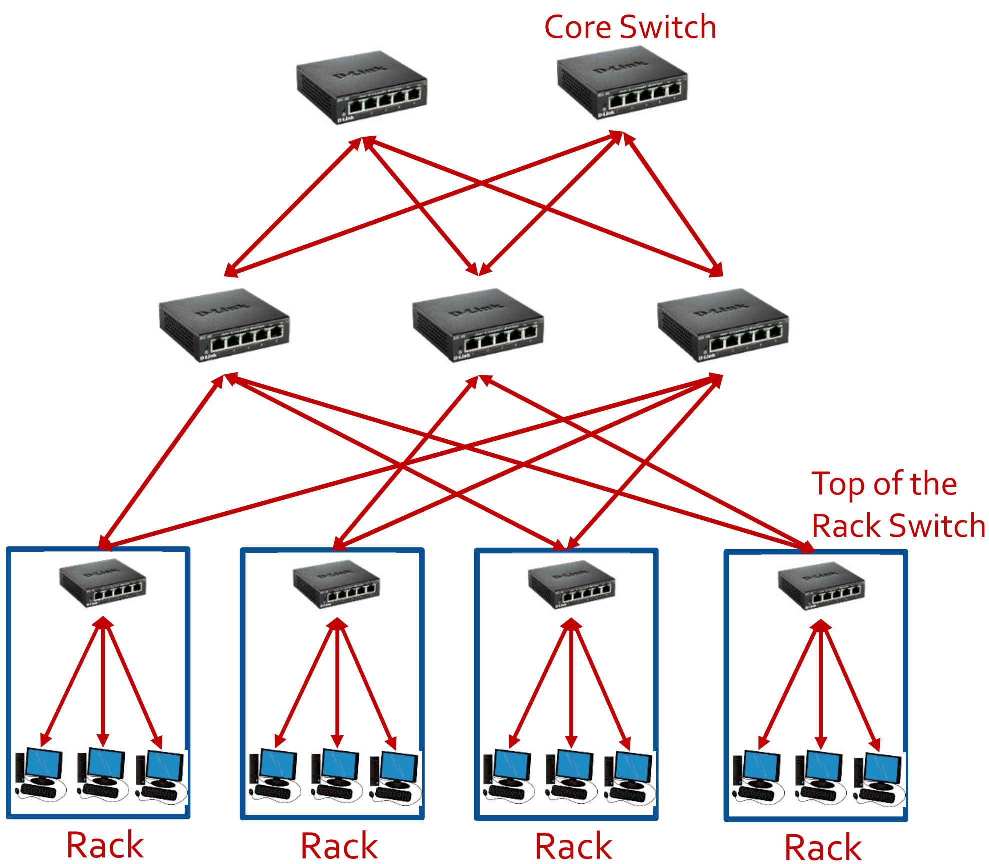


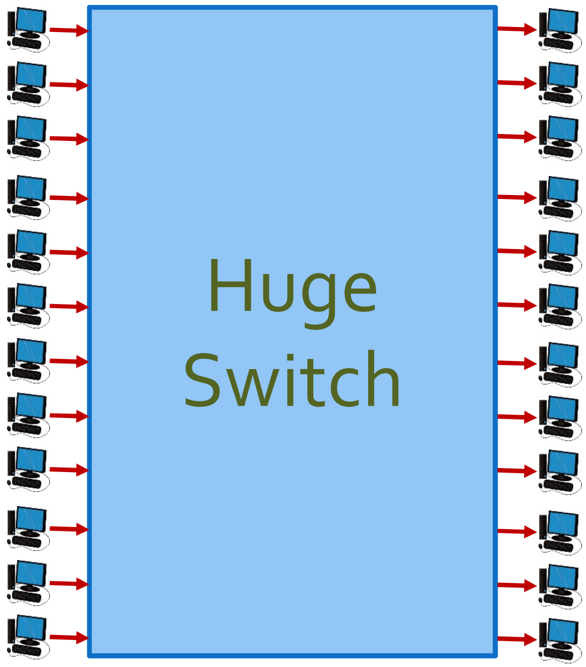
Transportation Systems

Call Centers

DATA CENTER NETWORKS

DATA CENTER NETWORKS

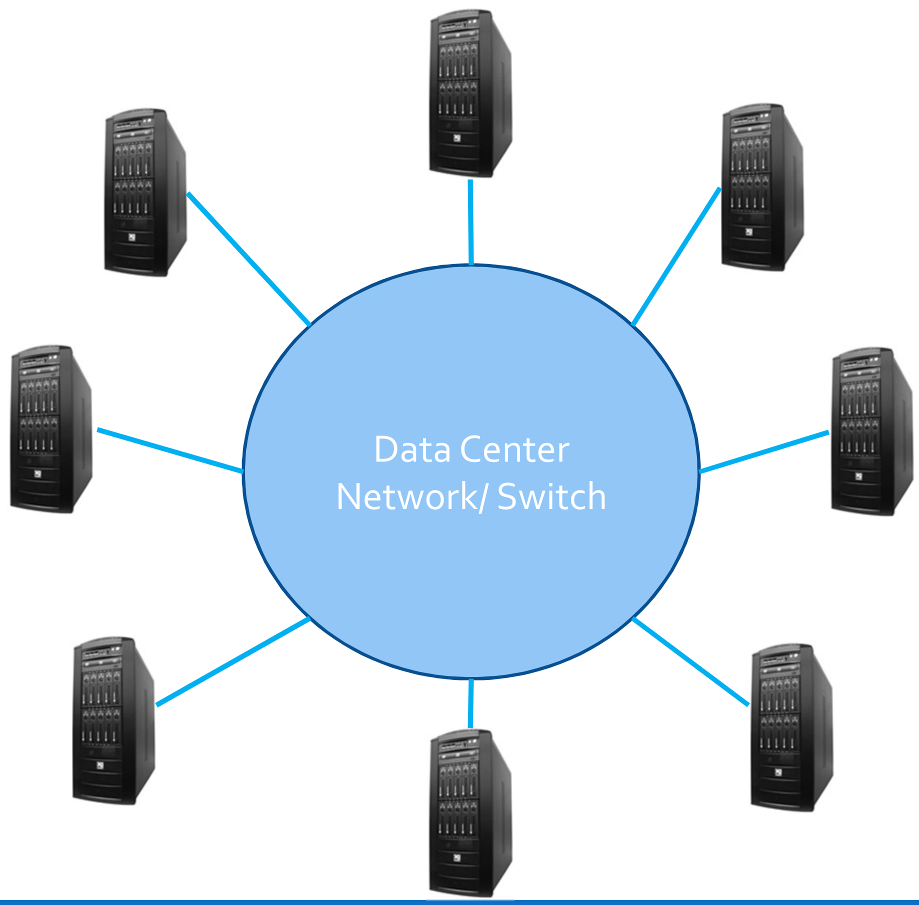




THE GRAND CHALLENGE

Are there low-complexity scheduling algorithms that maintain small packet delays, independent of the size of the network?

Yes – in heavy-traffic



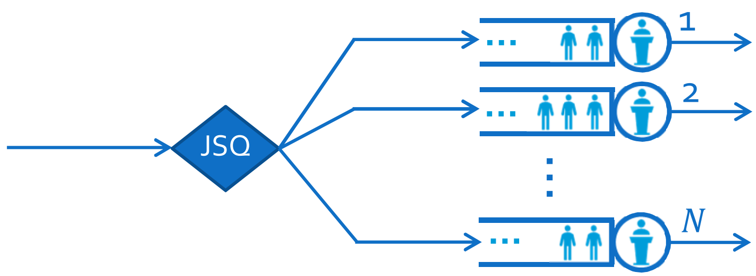
OUTLINE

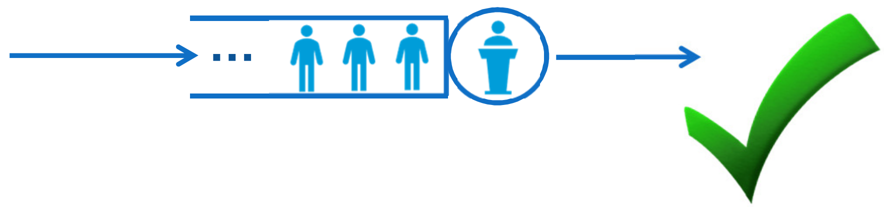
- Heavy-traffic analysis
- Drift method
  - Single server queue
  - Other queueing systems
- Concluding remarks

OUTLINE

- Heavy-traffic analysis
- Drift method
  - Single server queue
  - Other queueing systems
- Concluding remarks

HEAVY-TRAFFIC ANALYSIS

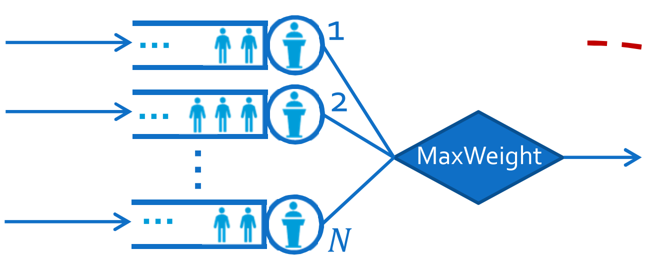


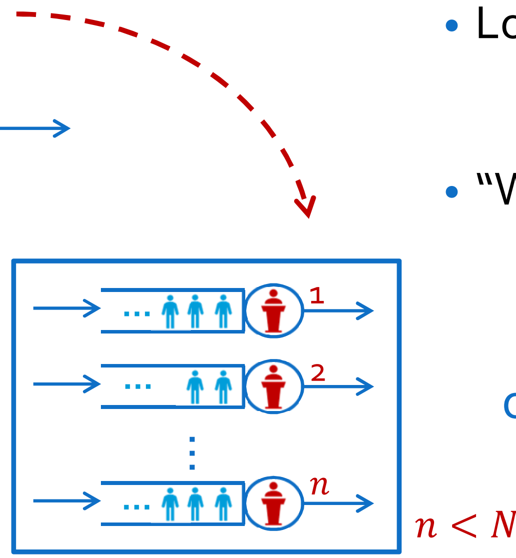


- Load system close to maximum capacity

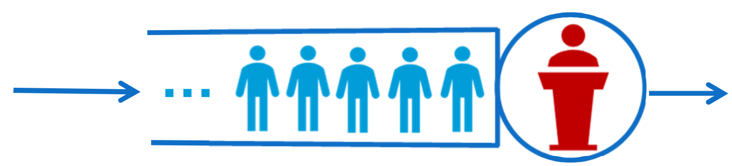
Arrival rate ≈ Service rate

- "Worst case scenario"





or



OUTLINE

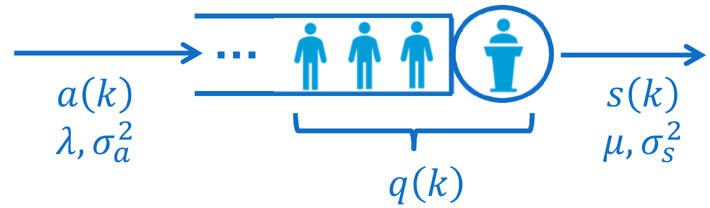
- Heavy-traffic analysis

- Drift method
  - Single server queue
  - Other queueing systems

- Concluding remarks

SINGLE SERVER QUEUE MODEL

- Discrete time model
- $q(k)$: # jobs in the system at the beginning of time slot k



- $a(k)$: # arrivals in time slot k
  - Mean λ and variance σ_a^2
- $s(k)$: **offered** service in time slot k
  - Mean μ and variance σ_s^2
- Both are i.i.d. sequences, independent of each other
- Stability: $\lambda < \mu$

- Dynamics of the queue

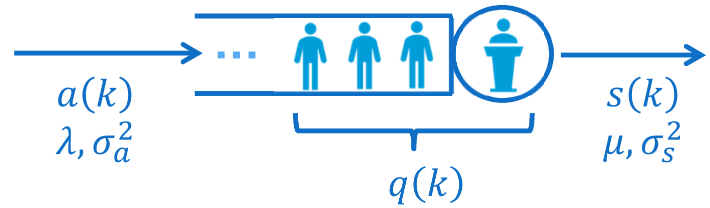
$$
\begin{aligned}
q(k+1) &= [q(k) + a(k) - s(k)]^+ \\
&= q(k) + a(k) - s(k) + u(k)
\end{aligned}
$$

Unused service ($u(k)$)

$$q(k+1)u(k) = 0$$

HEAVY TRAFFIC – DRIFT METHOD [Kingman, 1962]

- Load system close to its maximum capacity
  - $\lambda = \mu - \epsilon$, where $\epsilon > 0$
  - Limit as $\epsilon \downarrow 0$



$$q(k+1) = [q(k) + a(k) - s(k)]^+$$
$$= q(k) + a(k) - s(k) + u(k)$$

$$q(k+1)u(k) = 0$$

- In steady state ($k \to \infty$):

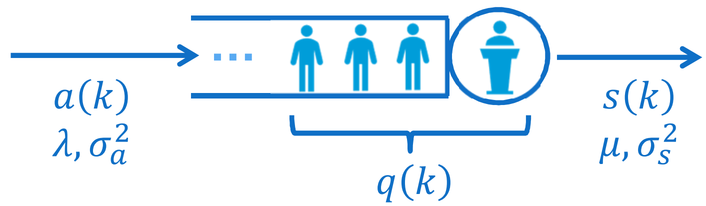
$$E[q^2(k+1)] = E[q^2(k)]$$

- Therefore,

$$\lim_{\epsilon \downarrow 0} E[\epsilon q] = \frac{\sigma_a^2 + \sigma_s^2}{2}$$

DRIFT METHOD [Eryilmaz, Srikant 13]

$\epsilon = \mu - \lambda$



- In steady state, set $E\left[V\left(q(k+1)\right)\right] = E\left[V\left(q(k)\right)\right]$

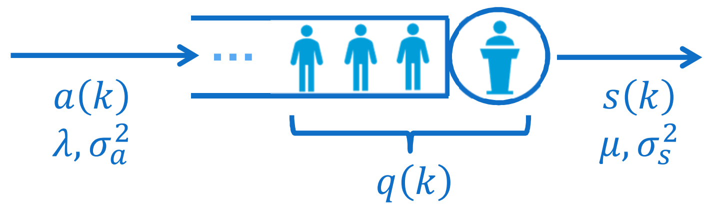
$V(q)$:	q^2	q^3	$\cdots$	q^{m+1}
Obtain:	$E[\epsilon q]$	$E[\epsilon^2 q^2]$	$\cdots$	$E[\epsilon^m q^m]$
	$\downarrow$	$\downarrow$		$\downarrow$
	$\frac{\sigma_a^2 + \sigma_s^2}{2}$	$2\left(\frac{\sigma_a^2 + \sigma_s^2}{2}\right)^2$		$m!\left(\frac{\sigma_a^2 + \sigma_s^2}{2}\right)^m$

Can we do this more efficiently?

$$\epsilon q \Rightarrow Expo\left(\frac{\sigma_a^2 + \sigma_s^2}{2}\right)$$

DRIFT METHOD [Eryilmaz, Srikant 13]

$\epsilon = \mu - \lambda$



- In steady state, set $E\left[V\left(q(k+1)\right)\right] = E\left[V\left(q(k)\right)\right]$

$V(q)$: $1 + \epsilon q + \frac{1}{2}\epsilon^2 q^2 + \frac{1}{3!}\epsilon^3 q^3 + \cdots + \frac{1}{(m+1)!}\epsilon^{m+1} q^{m+1} + \cdots = e^{\epsilon q}$

Obtain: $1 + E[\epsilon q] + \frac{1}{2}E[\epsilon^2 q^2] + \cdots + \frac{1}{m!}E[\epsilon^m q^m] + \cdots = E[e^{\epsilon q}]$

$E[\epsilon q] \to \frac{\sigma_a^2 + \sigma_s^2}{2}$, $\quad E[\epsilon^2 q^2] \to 2\left(\frac{\sigma_a^2 + \sigma_s^2}{2}\right)^2$, $\quad E[\epsilon^m q^m] \to m!\left(\frac{\sigma_a^2 + \sigma_s^2}{2}\right)^m$

$$\epsilon q \Rightarrow Expo\left(\frac{\sigma_a^2 + \sigma_s^2}{2}\right)$$

MOMENT GENERATING FUNCTION (MGF)

- Also known as Laplace transform
- **Definition**: MGF of r.v. X is $\mathrm{f}(\theta) = E\left[e^{\theta X}\right]$, where $\theta \in \mathbb{R}$ is a parameter
- The MGF completely determines the distribution of X
- Property:

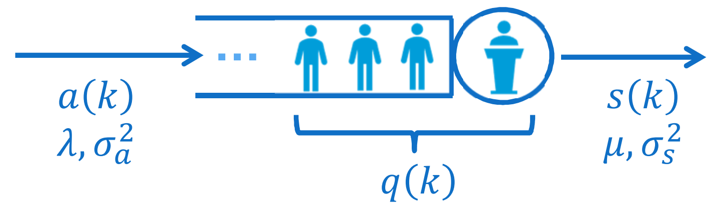
$$\frac{d}{d\theta} E\left[e^{\theta X}\right] = E\left[X e^{\theta X}\right] \quad \Rightarrow \quad \left.\frac{d}{d\theta} E\left[e^{\theta X}\right]\right|_{\theta=0} = E[X]$$

$$\frac{d^2}{d\theta^2} E\left[e^{\theta X}\right] = E\left[X^2 e^{\theta X}\right] \quad \Rightarrow \quad \left.\frac{d^2}{d\theta^2} E\left[e^{\theta X}\right]\right|_{\theta=0} = E[X^2]$$

$$\frac{d^m}{d\theta^m} E\left[e^{\theta X}\right] = E\left[X^m e^{\theta X}\right] \quad \Rightarrow \quad \left.\frac{d^m}{d\theta^m} E\left[e^{\theta X}\right]\right|_{\theta=0} = E[X^m]$$

DRIFT METHOD [Eryilmaz, Srikant 13]

$\epsilon = \mu - \lambda$



- In steady state, set $E\left[V\left(q(k+1)\right)\right] = E\left[V\left(q(k)\right)\right]$

$V(q)$: $$1 + \epsilon q + \frac{1}{2}\epsilon^2 q^2 + \frac{1}{3!}\epsilon^3 q^3 + \cdots + \frac{1}{(m+1)!}\epsilon^{m+1} q^{m+1} + \cdots = e^{\epsilon q}$$

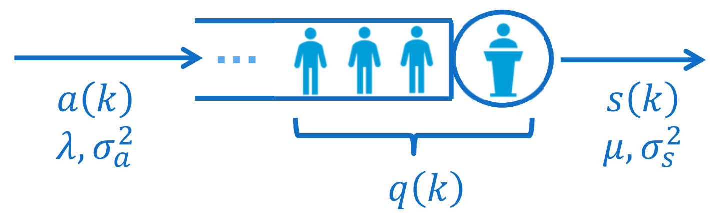
Obtain: $$1 + E[\epsilon q] + \frac{1}{2}E[\epsilon^2 q^2] + \cdots + \frac{1}{m!}E[\epsilon^m q^m] + \cdots = E[e^{\epsilon q}]$$

$$E[\epsilon q] \rightarrow \frac{\sigma_a^2 + \sigma_s^2}{2}, \quad E[\epsilon^2 q^2] \rightarrow 2\left(\frac{\sigma_a^2 + \sigma_s^2}{2}\right)^2, \quad E[\epsilon^m q^m] \rightarrow m!\left(\frac{\sigma_a^2 + \sigma_s^2}{2}\right)^m$$

$$\epsilon q \Rightarrow Expo\left(\frac{\sigma_a^2 + \sigma_s^2}{2}\right)$$

MGF METHOD [HL, Maguluri 19]

$\epsilon = \mu - \lambda$



- In steady-state: $E\left[e^{\theta\epsilon q(k+1)}\right] = E\left[e^{\theta\epsilon q(k)}\right]$

- Key Lemma:

$$\left(e^{\theta\epsilon q(k+1)} - 1\right)\left(e^{-\theta\epsilon u(k)} - 1\right) = 0$$

$$q(k+1) = q(k) + a(k) - s(k) + u(k)$$

$$q(k+1)u(k) = 0$$

- Rearranging terms

$$E\left[e^{\theta\epsilon q}\right] = \frac{1-E\left[e^{-\theta\epsilon u}\right]}{1-E\left[e^{\theta\epsilon(a-s)}\right]}$$

$\epsilon \downarrow 0$

$$\frac{1}{1-\theta\left(\frac{\sigma_a^2+\sigma_s^2}{2}\right)}$$

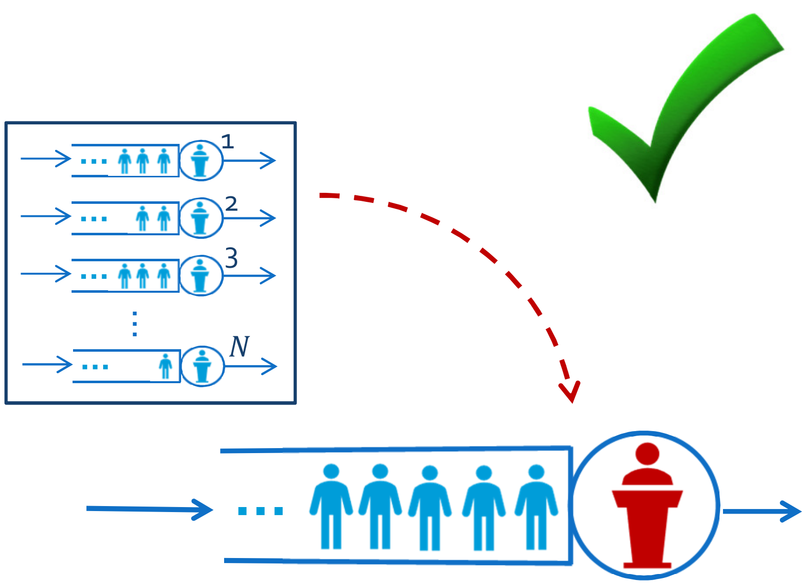
MGF of expo. r.v. with mean $\frac{\sigma_a^2+\sigma_s^2}{2}$

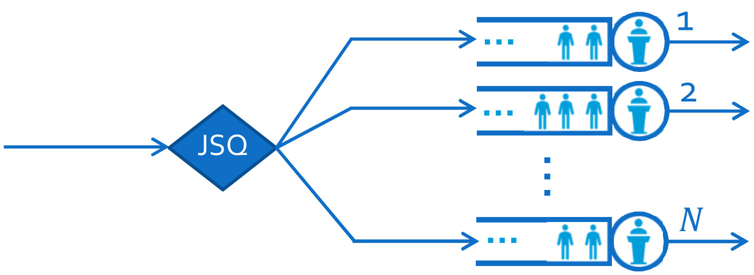
$$\epsilon q \Rightarrow Expo\left(\frac{\sigma_a^2+\sigma_s^2}{2}\right)$$

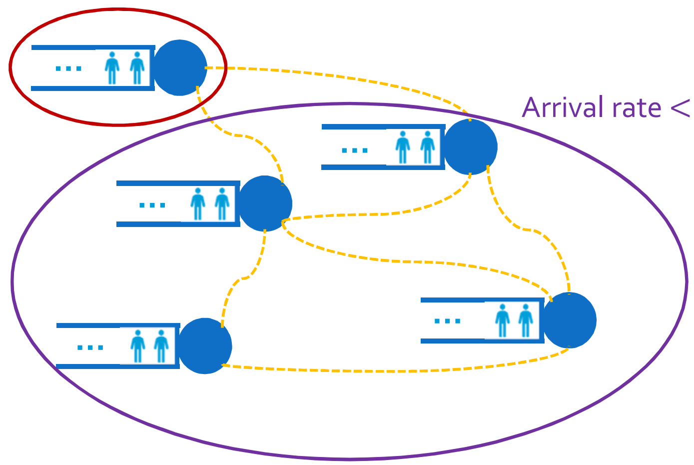
OTHER QUEUEING SYSTEMS?

- State Space Collapse (SSC): In heavy-traffic, the queueing systems behaves as if it lived in a lower dimensional state space

COMPLETE RESOURCE POOLING (CRP)



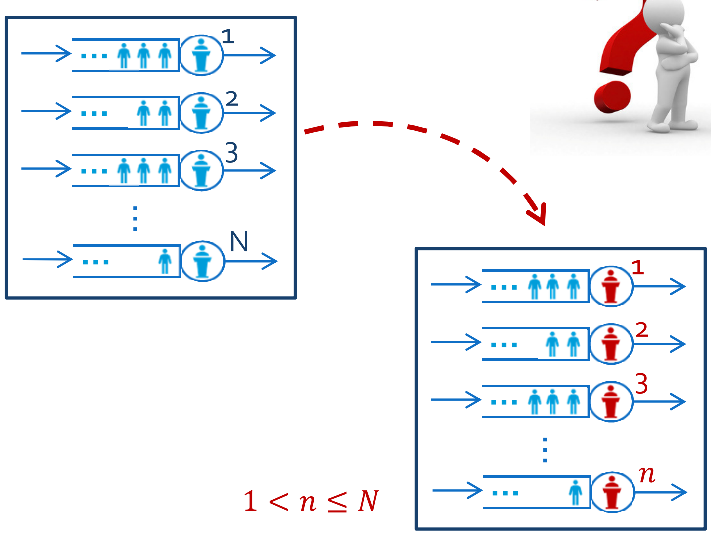


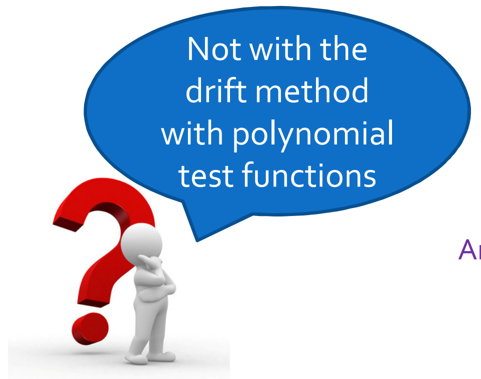


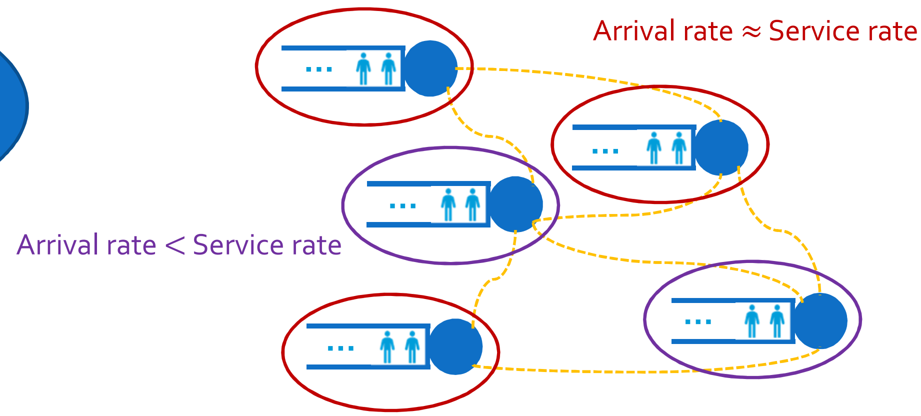
OTHER QUEUEING SYSTEMS? (CONT.)

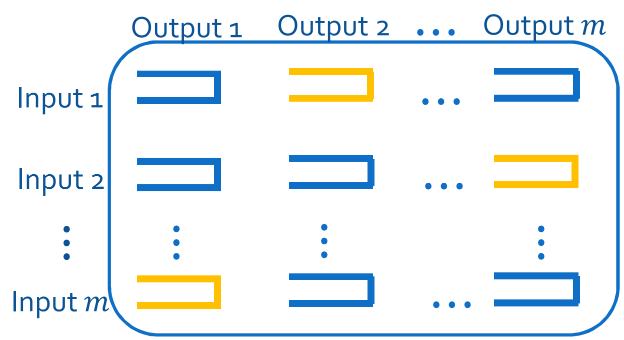
- State Space Collapse (SSC): In heavy-traffic, the queueing systems behaves as if it lived in a lower dimensional state space

NO CRP



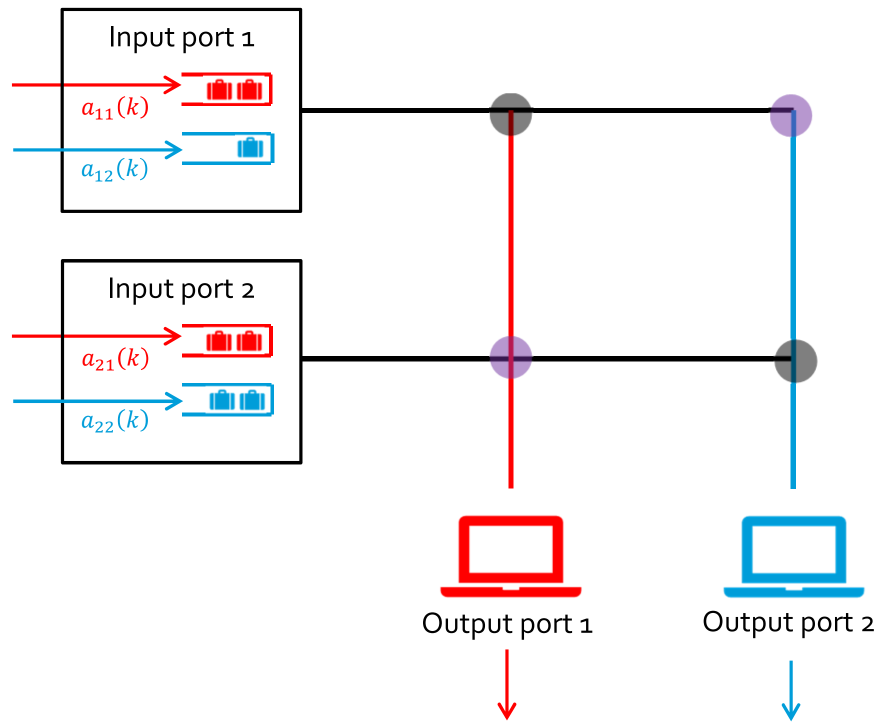






INPUT QUEUED SWITCH

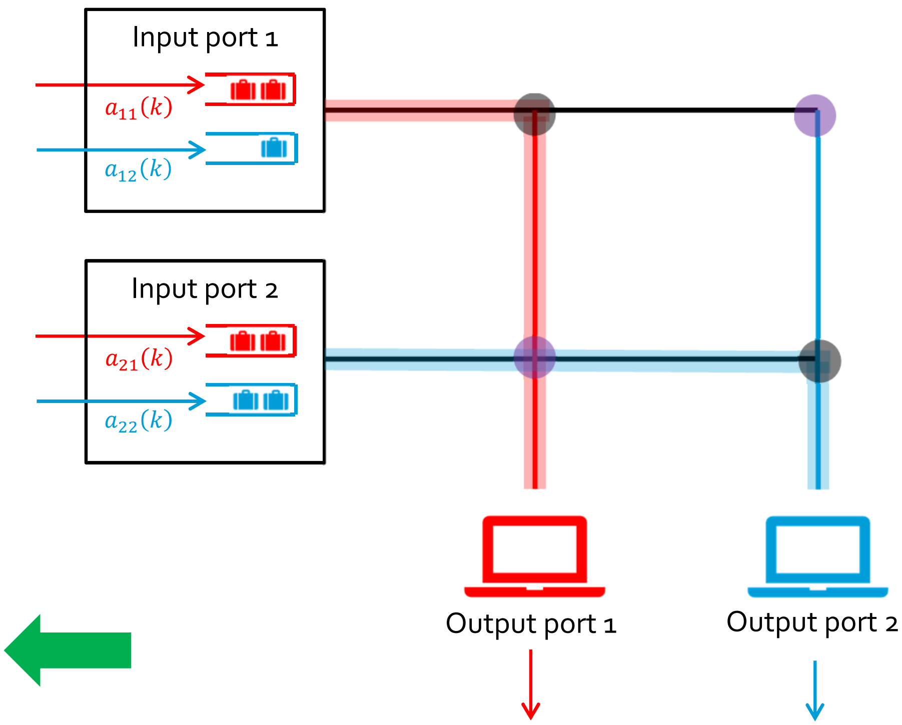
- Discrete time queueing system
- m input ports and m output ports
- Packets arrive to each input port with a predetermined output port
  - One queue for each input/output pair
- All packets are size 1
- All input ports can be connected to all output ports
- **Constraint**: Process at most one packet from each input port and at most one at each output port in each time slot



Which queues should we serve?

MAXWEIGHT ALGORITHM

- MaxWeight = Choose connection with maximum weight
- Weight = Total queue length of the connection

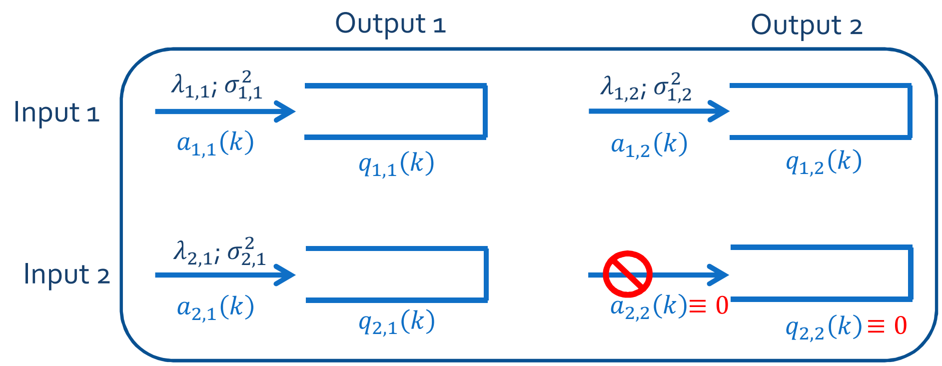


	Input 1	Input 2	Weight
Connection 1	Output 1	Output 2	2+2 = 4
Connection 2	Output 2	Output 1	1+2 = 3

NEW VIEW OF THE DRIFT METHOD

- Simplest no-CRP queueing system
  - 2x2 switch operating under MaxWeight
  - No arrivals to queue 2,2
  - Arrivals to other queues with mean $\lambda_{i,j}$ and variance $\sigma_{i,j}^2$
- Stability:
  - Arrival rate < Service rate
  - $\lambda_{1,1} + \lambda_{1,2} < 1$ & $\lambda_{1,1} + \lambda_{2,1} < 1$
- Heavy-traffic: $\epsilon > 0$
  - Arrival rate ≈Service rate
  - $\lambda_{1,1} = 1 - \lambda - \epsilon$ & $\lambda_{1,2} = \lambda_{2,1} = \lambda$

$\Rightarrow \lambda_{1,1} + \lambda_{1,2} = 1 - \epsilon$ & $\lambda_{1,1} + \lambda_{2,1} = 1 - \epsilon$



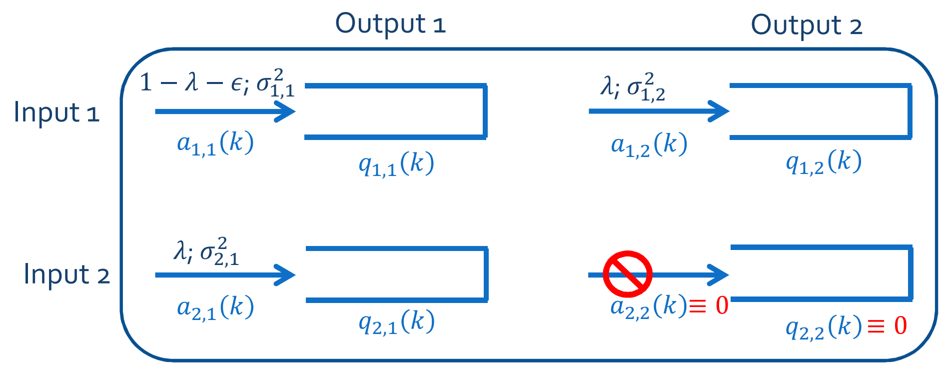
State Space Collapse [Maguluri et.al. 18]

2-dimensional
⇒No-CRP

$$\mathcal{K} = \left\{ \begin{bmatrix} x_{1,1} & x_{1,2} \\ x_{2,1} & 0 \end{bmatrix} : x_{1,1} = x_{1,2} + x_{2,1} \right\}$$

NEW VIEW OF THE DRIFT METHOD

- SSC:
  - $\boldsymbol{q}_{\parallel}$: Projection of $\boldsymbol{q}$ on $\mathcal{K}$
  - $E[\|\boldsymbol{q}\|^2] \approx E\left[\|\boldsymbol{q}_{\parallel}\|^2\right]$
  - $q_{\parallel 1,1} = q_{\parallel 1,2} + q_{\parallel 2,1}$



$$\mathcal{K} = \left\{\begin{bmatrix} x_{1,1} & x_{1,2} \\ x_{2,1} & 0 \end{bmatrix} : x_{1,1} = x_{1,2} + x_{2,1}\right\}$$

- Most general quadratic test function

$$V(\boldsymbol{q}) = \alpha_1 q_{\parallel 1,2}^2 + \alpha_2 q_{\parallel 2,1}^2 + \alpha_3 q_{\parallel 1,2} q_{\parallel 2,1}$$

$$V_1(\boldsymbol{q}) = q_{\parallel 1,2}^2$$
$$V_2(\boldsymbol{q}) = q_{\parallel 2,1}^2$$
$$V_3(\boldsymbol{q}) = q_{\parallel 1,2} q_{\parallel 2,1}$$

NEW VIEW OF THE DRIFT METHOD (cont.)

- Setting $E\left[V_i\left(\boldsymbol{q}(k+1)\right)\right] = E\left[V_i\left(\boldsymbol{q}(k)\right)\right]$ for $i = 1,2,3$:

(1) $$2\lim_{\epsilon\downarrow 0} \epsilon E\left[q_{1,2}\right] = \frac{\sigma_{1,1}^2 + 4\sigma_{1,2}^2 + \sigma_{2,1}^2}{3} - 2\lim_{\epsilon\downarrow 0} E\left[q_{1,2}^{+} u_{2,1}\right]$$

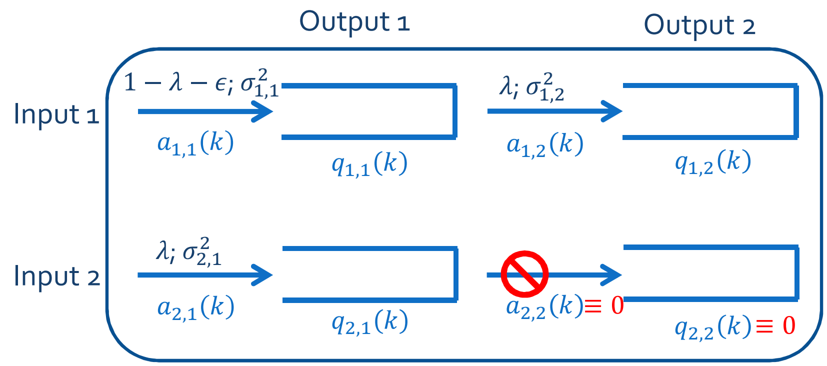
(2) $$2\lim_{\epsilon\downarrow 0} \epsilon E\left[q_{2,1}\right] = \frac{\sigma_{1,1}^2 + \sigma_{1,2}^2 + 4\sigma_{2,1}^2}{3} - 2\lim_{\epsilon\downarrow 0} E\left[q_{2,1}^{+} u_{1,2}\right]$$

(3) $$\lim_{\epsilon\downarrow 0} \epsilon E\left[q_{1,2} + q_{2,1}\right] = \frac{\sigma_{1,1}^2 - 2\sigma_{1,2}^2 - 2\sigma_{2,1}^2}{3} + 2\lim_{\epsilon\downarrow 0} E\left[q_{1,2}^{+} u_{2,1} + q_{2,1}^{+} u_{1,2}\right]$$

Theorem **[Maguluri et.al. 18]:**

$$\lim_{\epsilon\downarrow 0} \epsilon E\left[q_{1,1} + q_{1,2} + q_{2,1}\right] = \frac{2}{3}\left(\sigma_{1,1}^2 + \sigma_{1,2}^2 + \sigma_{2,1}^2\right)$$

Proof: (1)+(2)+(3) and rearrange terms



OTHER LINEAR COMBINATIONS?

4 unknowns
3 equations !

Need more equations !

CONCLUDING REMARKS

COMPLETE RESOURCE POOLING (CRP)

- Queueing system behaves as single server queue
- MGF method to obtain distribution of ϵq in heavy-traffic
- It's all about handling unused service

$$q(k+1)u(k) = 0$$

$$\left(e^{\theta\epsilon q(k+1)} - 1\right)\left(e^{-\theta\epsilon u(k)} - 1\right) = 0$$

CRP CONDITION NOT SATISFIED

- Queueing system **does not** behave as single server queue
- Only first moment can be obtained from drift method with polynomial test functions
- Future work: Other moments and distribution

OPTIMAL RESOURCE ALLOCATION IN DATA CENTER NETWORKS: DRIFT METHOD AND TRANSFORM TECHNIQUES

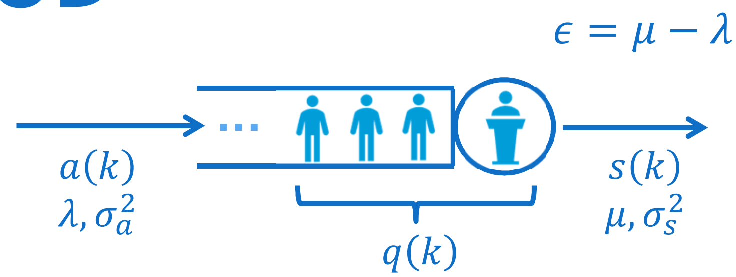
Daniela Hurtado-Lange, Siva Theja Maguluri

Industrial and Systems Engineering

Georgia Institute of Technology

September 2019

DRIFT METHOD



- Drift method [Eryilmaz, Srikant 13]:

$V(q)$: q^2 $\quad q^3 \quad \cdots \quad q^{m+1}$

Obtain: $E[\epsilon q]$ $\quad E[\epsilon^2 q^2] \quad \cdots \quad E[\epsilon^m q^m]$

$$E[\epsilon q] \downarrow \frac{\sigma_a^2 + \sigma_s^2}{2}$$

$$E[\epsilon^2 q^2] \downarrow 2\left(\frac{\sigma_a^2 + \sigma_s^2}{2}\right)^2$$

$$E[\epsilon^m q^m] \downarrow m!\left(\frac{\sigma_a^2 + \sigma_s^2}{2}\right)^m$$

$$\epsilon q \Rightarrow Expo\left(\frac{\sigma_a^2 + \sigma_s^2}{2}\right)$$