Pass-and-Swap Queues

Joint work with Jan-Pieter Dorsman (UvA)

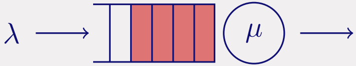
Céline Comte — c.m.comte@tue.nl
Eindhoven University of Technology

2021 INFORMS Annual Meeting — Session “Recent Advances in Load Balancing”

The M/M/1 queue

Model

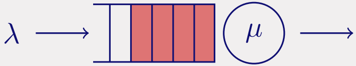
- Jobs arrive according to a Poisson process with rate λ.
- Service times are i.i.d. exponentially distributed with rate μ.
- A single server.



The M/M/1 queue

Model

- Jobs arrive according to a Poisson process with rate λ.
- Service times are i.i.d. exponentially distributed with rate μ.
- A single server.



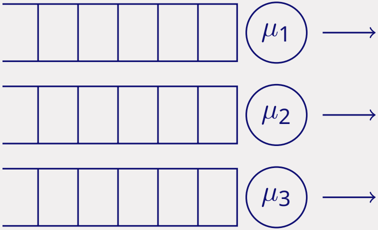
Analysis

- Markov (birth-and-death) process.
- Stationary distribution: $\pi(n) = \left(1 - \frac{\lambda}{\mu}\right)\left(\frac{\lambda}{\mu}\right)^n$.
- Proof: (partial) balance equations + normalization equation.

Order-independent queues

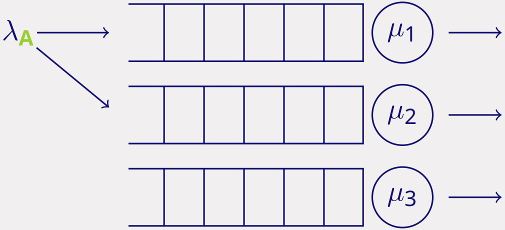
Order-independent queues

- **Redundancy cancel-on-complete** (Gardner et al., 2016)



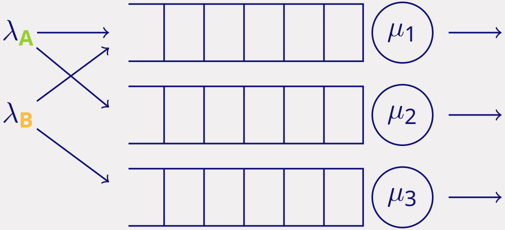
Order-independent queues

- **Redundancy cancel-on-complete** (Gardner et al., 2016)



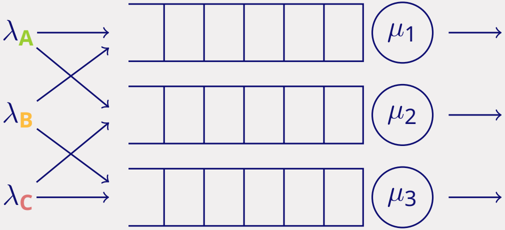
Order-independent queues

- **Redundancy cancel-on-complete** (Gardner et al., 2016)



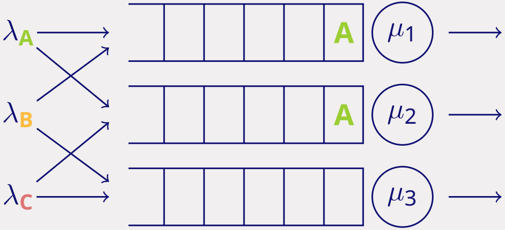
Order-independent queues

- **Redundancy cancel-on-complete** (Gardner et al., 2016)



Order-independent queues

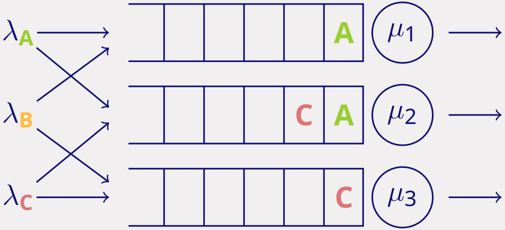
- **Redundancy cancel-on-complete** (Gardner et al., 2016)



TU/e

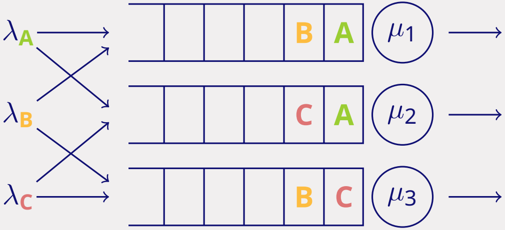
Order-independent queues

- **Redundancy cancel-on-complete** (Gardner et al., 2016)



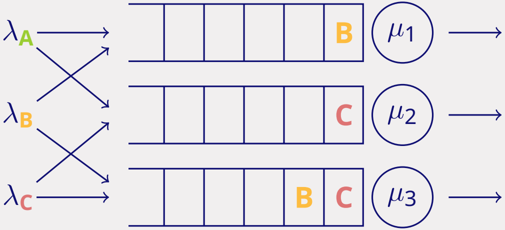
Order-independent queues

- **Redundancy cancel-on-complete** (Gardner et al., 2016)



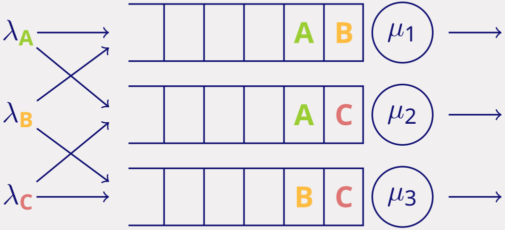
Order-independent queues

- **Redundancy cancel-on-complete** (Gardner et al., 2016)



Order-independent queues

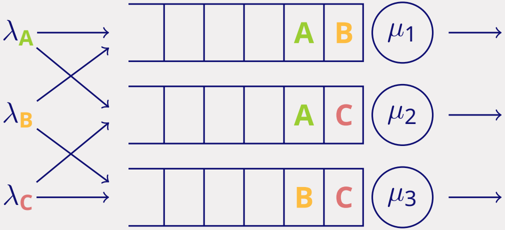
- **Redundancy cancel-on-complete** (Gardner et al., 2016)



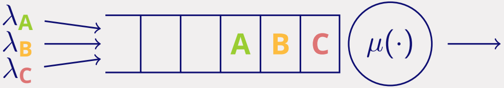
TU/e

Order-independent queues

- **Redundancy cancel-on-complete** (Gardner et al., 2016)



- **Order-independent (OI) queue**
  (Berezner et al., 1995)
  (Bonald and Comte, 2017)



Order-independent queues

- Product-form stationary distribution (Gardner et al., 2016):

$$\pi(c_1, c_2, \ldots, c_n) = \pi(\varnothing) \prod_{p=1}^{n} \frac{\lambda_{c_p}}{\mu(c_1, \ldots, c_p)}.$$

Order-independent queues

- Product-form stationary distribution (Gardner et al., 2016):

$$\pi(c_1, c_2, \ldots, c_n) = \pi(\varnothing) \prod_{p=1}^{n} \frac{\lambda_{c_p}}{\mu(c_1, \ldots, c_p)}.$$

- Proof: (partial) balance equations + normalization equation.

Order-independent queues

- Product-form stationary distribution (Gardner et al., 2016):

$$\pi(c_1, c_2, \dots, c_n) = \pi(\emptyset) \prod_{p=1}^{n} \frac{\lambda_{c_p}}{\mu(c_1, \dots, c_p)}.$$

- Proof: (partial) balance equations + normalization equation.
- Why study product-form queues?
  - Rich in applications.
  - Exact performance analysis is not completely hopeless.

Current state of the art

- Fundamental contributions by Kelly and Whittle in the 1970's and 1980's.

Current state of the art

- Fundamental contributions by Kelly and Whittle in the 1970’s and 1980’s.
- Study of product-form queues in different contexts: redundancy scheduling, matching systems, manufacturing systems, etc.

Current state of the art

- Fundamental contributions by Kelly and Whittle in the 1970’s and 1980’s.
- Study of product-form queues in different contexts: redundancy scheduling, matching systems, manufacturing systems, etc.
- Q: Can we build a model that captures all product-form queues out there?

Current state of the art

- Fundamental contributions by Kelly and Whittle in the 1970's and 1980's.
- Study of product-form queues in different contexts: redundancy scheduling, matching systems, manufacturing systems, etc.
- Q: Can we build a model that captures all product-form queues out there?
- A: Several frameworks have been constructed recently:
  - Adan, Kleiner, Righter, Weiss (2018)
  - Gardner, Righter (2020)
  - Ayesta, Bodas, Dorsman, Verloop (2021)

Current state of the art

- Fundamental contributions by Kelly and Whittle in the 1970's and 1980's.
- Study of product-form queues in different contexts: redundancy scheduling, matching systems, manufacturing systems, etc.
- Q: Can we build a model that captures all product-form queues out there?
- A: Several frameworks have been constructed recently:
  - Adan, Kleiner, Righter, Weiss (2018)
  - Gardner, Righter (2020)
  - Ayesta, Bodas, Dorsman, Verloop (2021)
- A: But, still, new product-form queues keep appearing, not captured by these frameworks, such as the pass-and-swap queue.

Definition

Pass-and-swap (P&S) queues

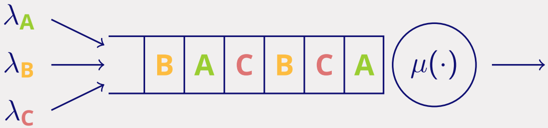
- are an extension of OI queues,
- add a whole new dimension to product-form queues: intra-queue routing,
- have many applications.

Definition

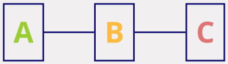
Pass-and-swap (P&S) queues

- are an extension of OI queues,
- add a whole new dimension to product-form queues: intra-queue routing,
- have many applications.

OI queue



Swapping graph

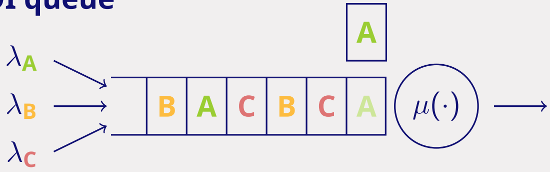


Definition

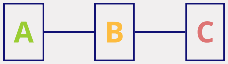
Pass-and-swap (P&S) queues

- are an extension of OI queues,
- add a whole new dimension to product-form queues: intra-queue routing,
- have many applications.

OI queue



Swapping graph

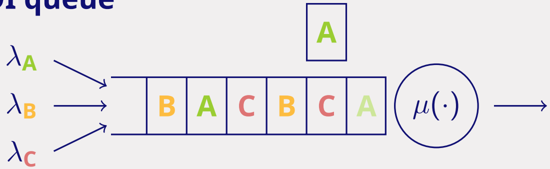


Definition

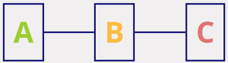
Pass-and-swap (P&S) queues

- are an extension of OI queues,
- add a whole new dimension to product-form queues: intra-queue routing,
- have many applications.

OI queue



Swapping graph

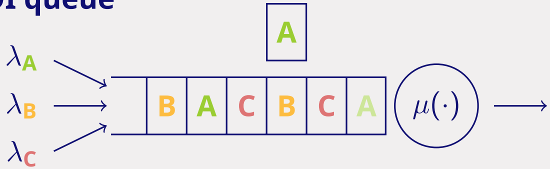


Definition

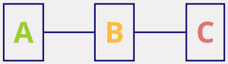
Pass-and-swap (P&S) queues

- are an extension of OI queues,
- add a whole new dimension to product-form queues: intra-queue routing,
- have many applications.

OI queue



Swapping graph

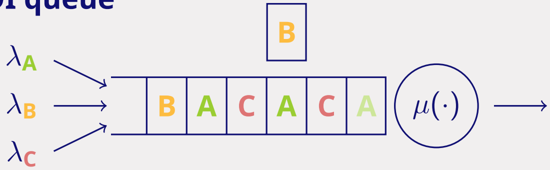


Definition

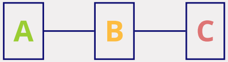
Pass-and-swap (P&S) queues

- are an extension of OI queues,
- add a whole new dimension to product-form queues: intra-queue routing,
- have many applications.

OI queue



Swapping graph

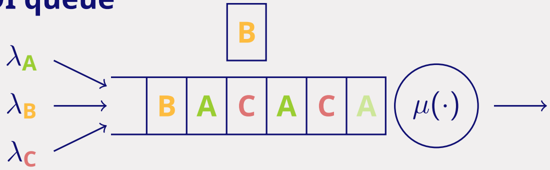


Definition

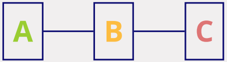
Pass-and-swap (P&S) queues

- are an extension of OI queues,
- add a whole new dimension to product-form queues: intra-queue routing,
- have many applications.

OI queue



Swapping graph

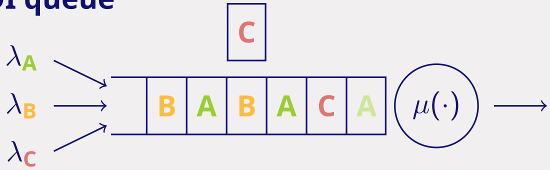


Definition

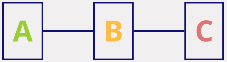
Pass-and-swap (P&S) queues

- are an extension of OI queues,
- add a whole new dimension to product-form queues: intra-queue routing,
- have many applications.

OI queue



Swapping graph

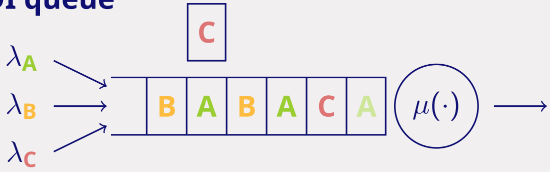


Definition

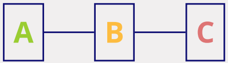
Pass-and-swap (P&S) queues

- are an extension of OI queues,
- add a whole new dimension to product-form queues: intra-queue routing,
- have many applications.

OI queue



Swapping graph

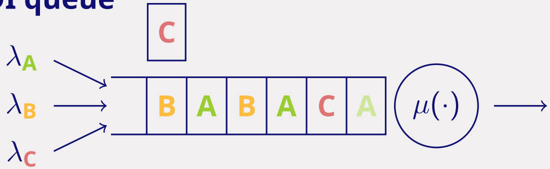


Definition

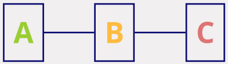
Pass-and-swap (P&S) queues

- are an extension of OI queues,
- add a whole new dimension to product-form queues: intra-queue routing,
- have many applications.

OI queue



Swapping graph

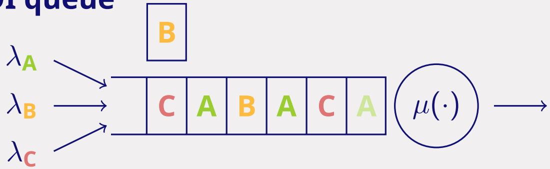


Definition

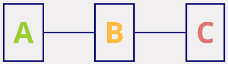
Pass-and-swap (P&S) queues

- are an extension of OI queues,
- add a whole new dimension to product-form queues: intra-queue routing,
- have many applications.

OI queue



Swapping graph

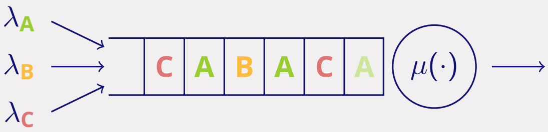


Definition

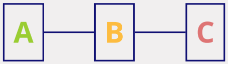
Pass-and-swap (P&S) queues

- are an extension of OI queues,
- add a whole new dimension to product-form queues: intra-queue routing,
- have many applications.

OI queue



Swapping graph

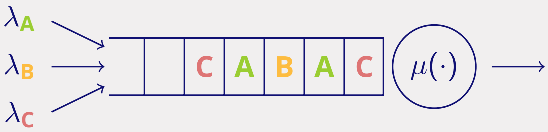


Definition

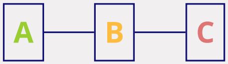
Pass-and-swap (P&S) queues

- are an extension of OI queues,
- add a whole new dimension to product-form queues: intra-queue routing,
- have many applications.

OI queue



Swapping graph

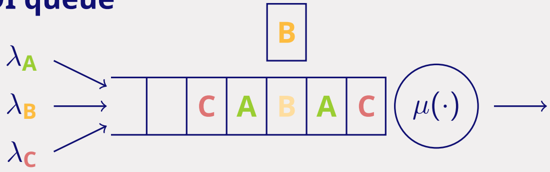


Definition

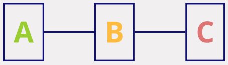
Pass-and-swap (P&S) queues

- are an extension of OI queues,
- add a whole new dimension to product-form queues: intra-queue routing,
- have many applications.

OI queue



Swapping graph

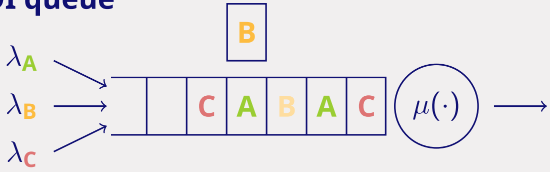


Definition

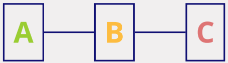
Pass-and-swap (P&S) queues

- are an extension of OI queues,
- add a whole new dimension to product-form queues: intra-queue routing,
- have many applications.

OI queue



Swapping graph

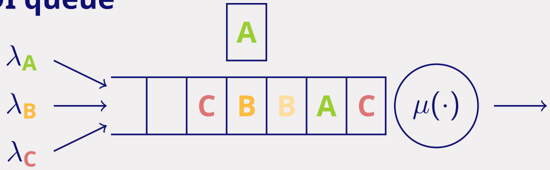


Definition

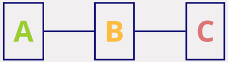
Pass-and-swap (P&S) queues

- are an extension of OI queues,
- add a whole new dimension to product-form queues: intra-queue routing,
- have many applications.

OI queue



Swapping graph

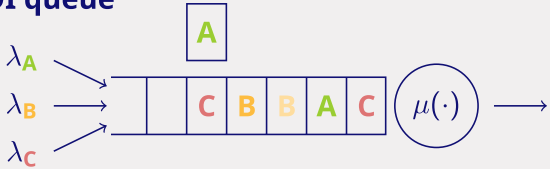


Definition

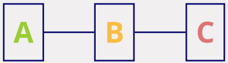
Pass-and-swap (P&S) queues

- are an extension of OI queues,
- add a whole new dimension to product-form queues: intra-queue routing,
- have many applications.

OI queue



Swapping graph

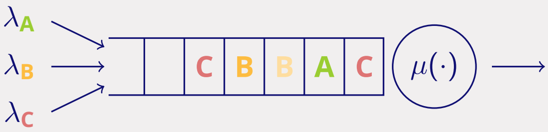


Definition

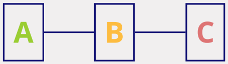
Pass-and-swap (P&S) queues

- are an extension of OI queues,
- add a whole new dimension to product-form queues: intra-queue routing,
- have many applications.

OI queue



Swapping graph

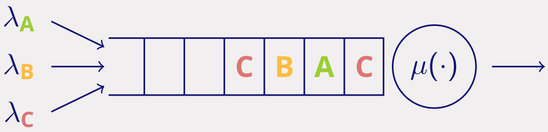


Definition

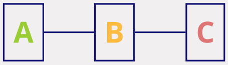
Pass-and-swap (P&S) queues

- are an extension of OI queues,
- add a whole new dimension to product-form queues: intra-queue routing,
- have many applications.

OI queue



Swapping graph



Product-form stationary distribution

- Stationary distribution: *exactly* the same as OI queues!

$$\pi(c_1, c_2, \dots, c_n) = \pi(\emptyset) \prod_{p=1}^{n} \frac{\lambda_{c_p}}{\mu(c_1, \dots, c_p)}.$$

TU/e

Product-form stationary distribution

- Stationary distribution: *exactly* the same as OI queues!

$$\pi(c_1, c_2, \ldots, c_n) = \pi(\varnothing) \prod_{p=1}^{n} \frac{\lambda_{c_p}}{\mu(c_1, \ldots, c_p)}.$$

- Proof: (partial) balance equations + normalization equation.

TU/e

Product-form stationary distribution

- Stationary distribution: *exactly* the same as OI queues!

$$\pi(c_1, c_2, \ldots, c_n) = \pi(\emptyset) \prod_{p=1}^{n} \frac{\lambda_{c_p}}{\mu(c_1, \ldots, c_p)}.$$

- Proof: (partial) balance equations + normalization equation.
- Hence, the P&S queue is a product-form queue.

Product-form stationary distribution

- Stationary distribution: *exactly* the same as OI queues!

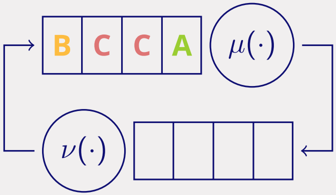
$$\pi(c_1, c_2, \ldots, c_n) = \pi(\varnothing) \prod_{p=1}^{n} \frac{\lambda_{c_p}}{\mu(c_1, \ldots, c_p)}.$$

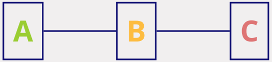
- Proof: (partial) balance equations + normalization equation.
- Hence, the P&S queue is a product-form queue.

- We also prove a simple stability condition (also valid for OI queues).

Closed network of P&S queues

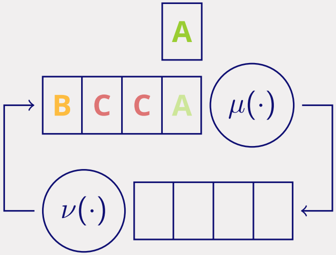
- Tandem network of two P&S queues with the same swapping graph:

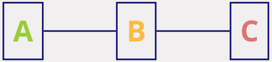




Closed network of P&S queues

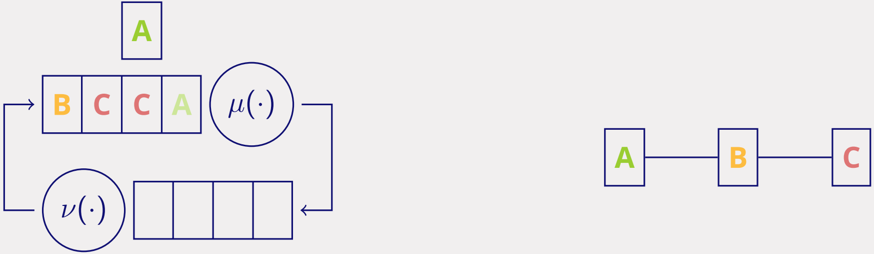
- Tandem network of two P&S queues with the same swapping graph:





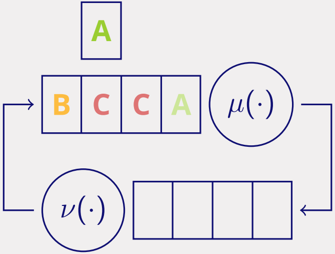
Closed network of P&S queues

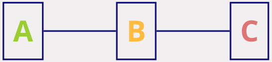
- Tandem network of two P&S queues with the same swapping graph:



Closed network of P&S queues

- Tandem network of two P&S queues with the same swapping graph:

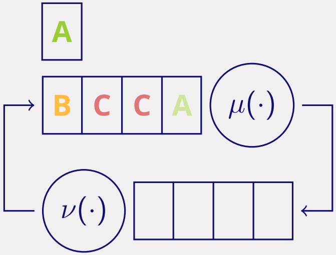


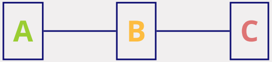


TU/e

Closed network of P&S queues

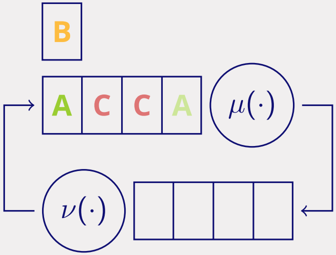
- Tandem network of two P&S queues with the same swapping graph:

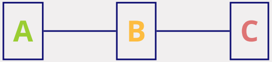




Closed network of P&S queues

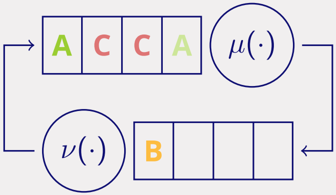
- Tandem network of two P&S queues with the same swapping graph:

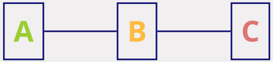




Closed network of P&S queues

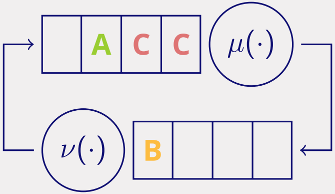
- Tandem network of two P&S queues with the same swapping graph:

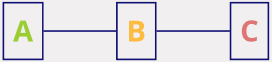




Closed network of P&S queues

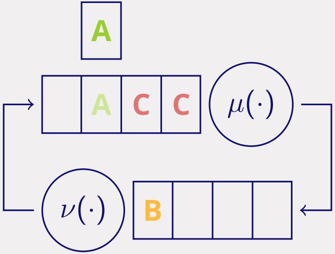
- Tandem network of two P&S queues with the same swapping graph:

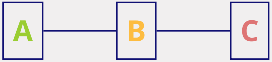




Closed network of P&S queues

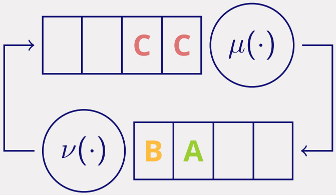
- Tandem network of two P&S queues with the same swapping graph:

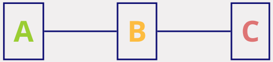




Closed network of P&S queues

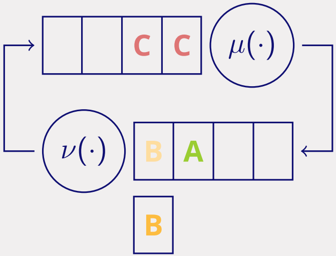
- Tandem network of two P&S queues with the same swapping graph:

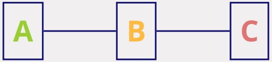




Closed network of P&S queues

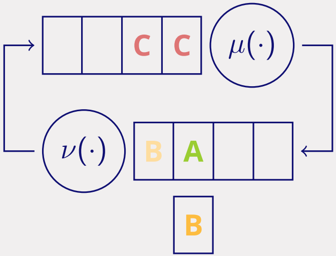
- Tandem network of two P&S queues with the same swapping graph:

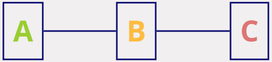




Closed network of P&S queues

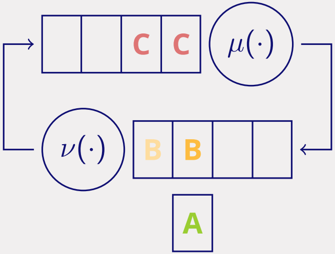
- Tandem network of two P&S queues with the same swapping graph:

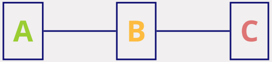




Closed network of P&S queues

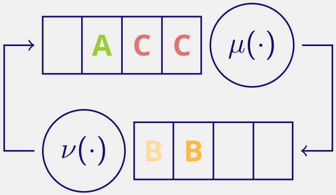
- Tandem network of two P&S queues with the same swapping graph:

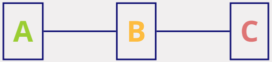




Closed network of P&S queues

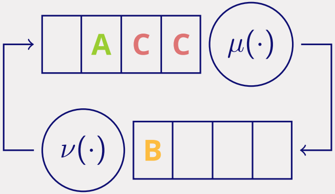
- Tandem network of two P&S queues with the same swapping graph:

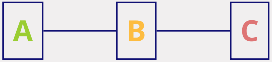




Closed network of P&S queues

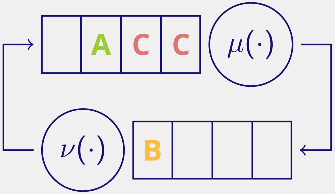
- Tandem network of two P&S queues with the same swapping graph:

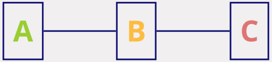




Closed network of P&S queues

- Tandem network of two P&S queues with the same swapping graph:

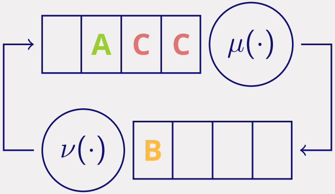


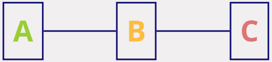


- B always last in the upper queue, first in the lower queue: *placement* order.

Closed network of P&S queues

- Tandem network of two P&S queues with the same swapping graph:

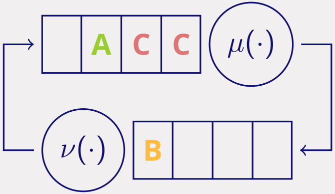




- B always last in the upper queue, first in the lower queue: *placement* order.
- The stationary distribution again has a product form (on a restricted space)!

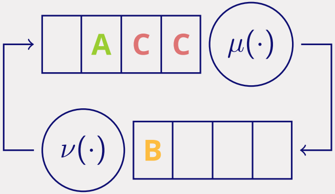
From two queues to one queue

- The lower queue models “state-dependent arrivals” to the upper queue.



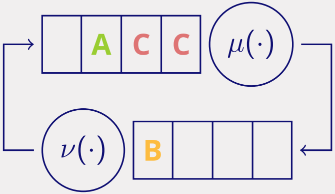
From two queues to one queue

- The lower queue models “state-dependent arrivals” to the upper queue.
- If the two queues are simple ·/M/1 queues, the upper queue can be seen as an M/M/1 queue with blocking.



From two queues to one queue

- The lower queue models “state-dependent arrivals” to the upper queue.
- If the two queues are simple ·/M/1 queues, the upper queue can be seen as an M/M/1 queue with blocking.
- This is very powerful:
  - Redundancy cancel-on-start and cancel-on-commit.
  - Hierarchical load-distribution algorithms.



Conclusion

Take away

- P&S queues broaden the family of product-form queues by allowing for intra-queue routing.
- Networks of P&S queues also have a product form.
- Paves the way for performance analysis of other algorithms.

Conclusion

Take away

- P&S queues broaden the family of product-form queues by allowing for intra-queue routing.
- Networks of P&S queues also have a product form.
- Paves the way for performance analysis of other algorithms.

Future works

- How big is the family of product-form queues?
- Are there other routing mechanisms that lead to a product form?
- Can we find other applications of P&S queues?