Resource pooling, proportional fairness and product form

Frank Kelly
University of Cambridge
www.statslab.cam.ac.uk
(includes work with Laurent Massoulié,
Neil Walton, Ruth Williams)

SIGMETRICS/Performance 2009,
Seattle

Outline

- The processor sharing queue
- Sharing in networks – proportional fairness
- A related queueing network – product form
- Heavy traffic for a flow model – proportional fairness *and* product form

Processor sharing discipline

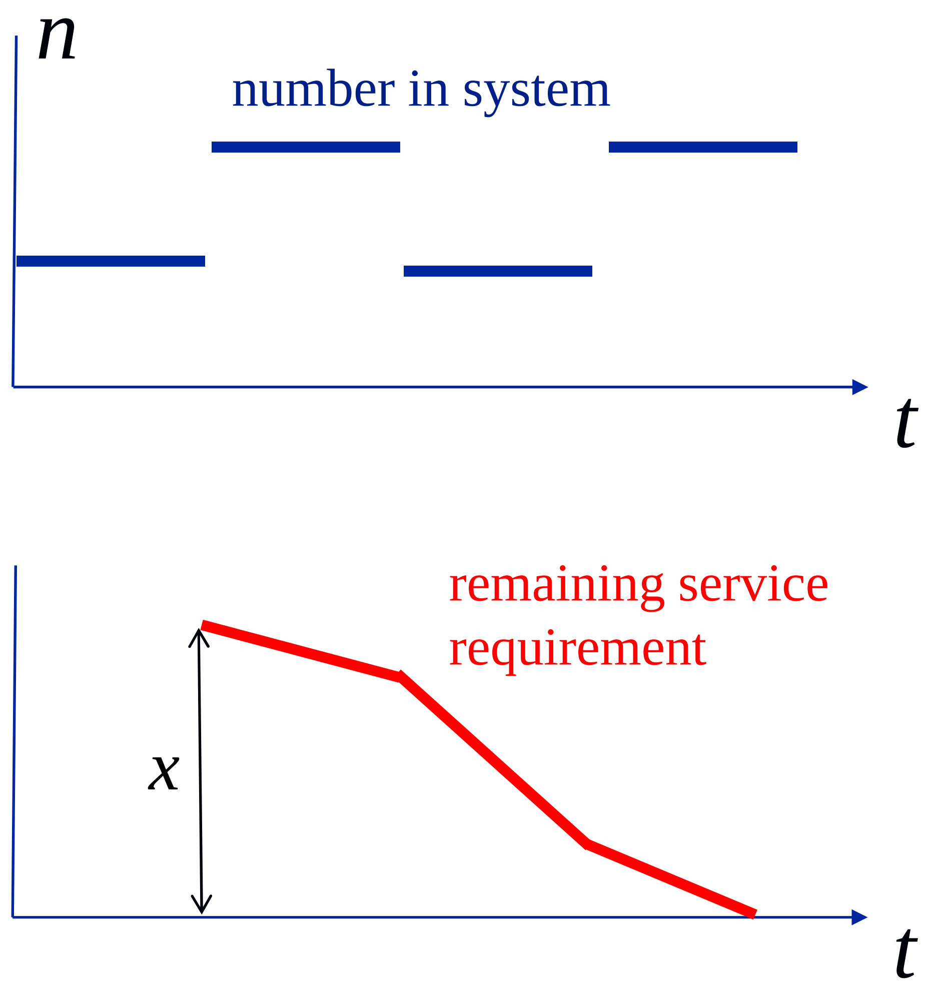
Kleinrock, 1967, 1976; Boxma tutorial, informs 2005

- Often attractive in practice, since gives
  - rapid service for short jobs
  - the appearance of a processor continuously available (albeit of varying capacity)
- Tractable analytically – a symmetric discipline. E.g. for M/G/1 PS

$$E\left[\text{sojourn time, } S \mid \text{job size, } x\right] = \frac{x}{C - \rho}$$

(similar tractability for LCFS, Erlang loss system, networks of symmetric queues)

The M/G/1 processor sharing queue



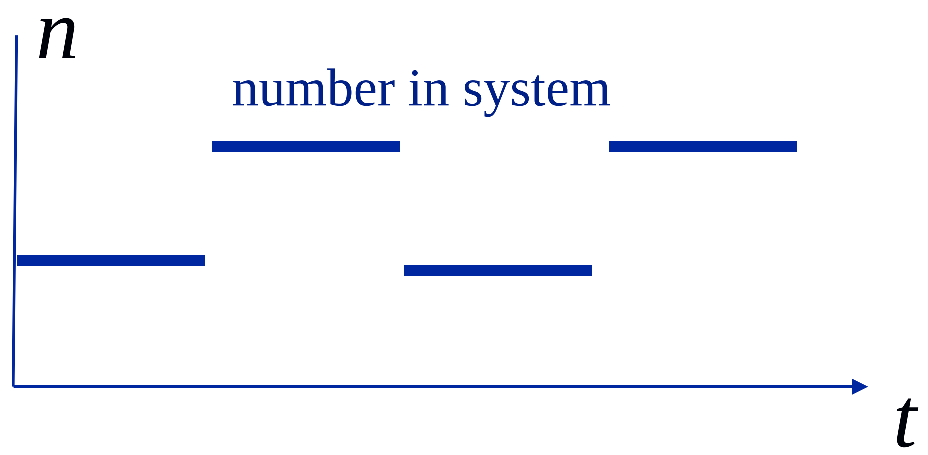
The M/G/1 processor sharing queue

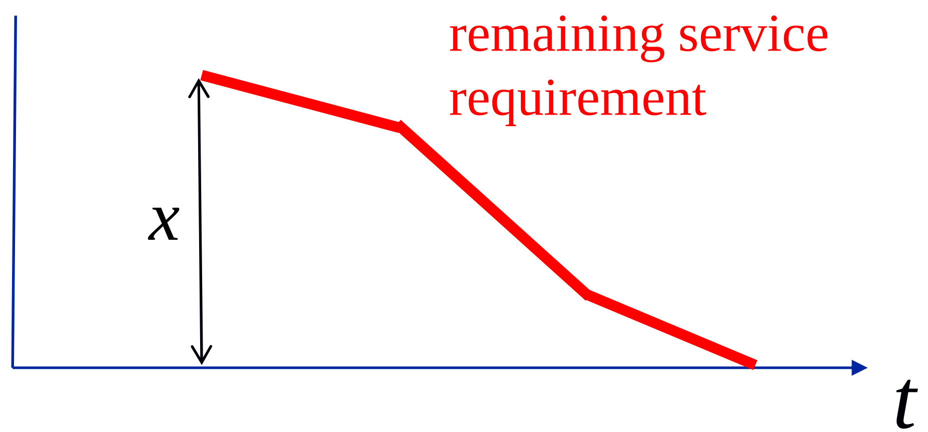
$$[S \mid x] \cong \frac{x}{C-\rho} + o(1/x)$$

if x is large;

$$[S \mid x] \cong x.\frac{n+1}{C} + o(x)$$

if x is small, where n is a geometric random variable.





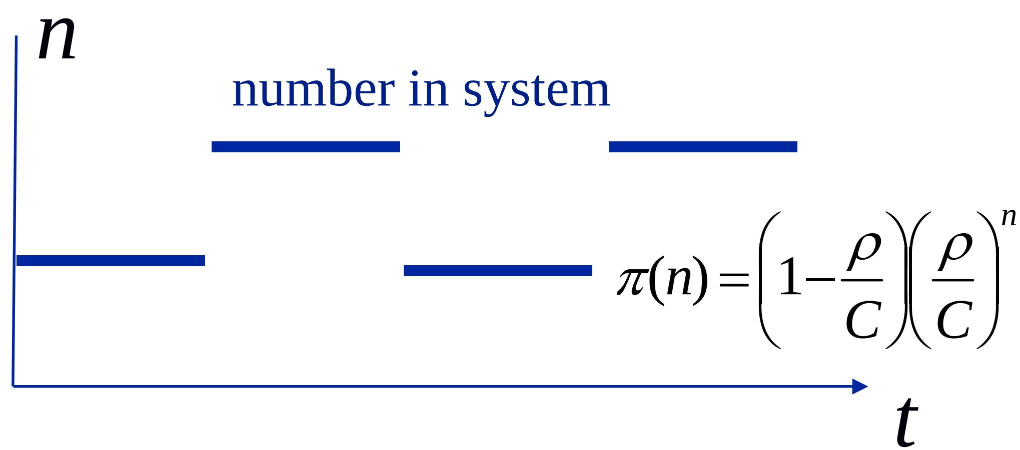
The M/G/1 processor sharing queue

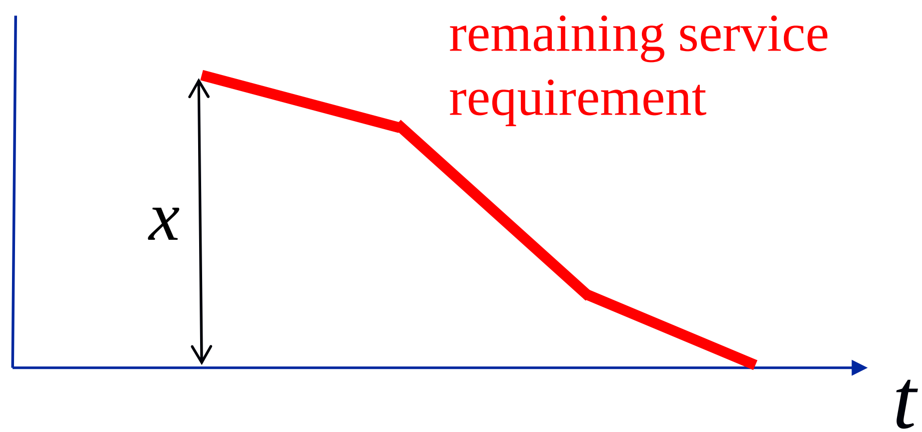
$$[S \mid x] \cong \frac{x}{C-\rho} + o(1/x)$$

if x is large;

$$[S \mid x] \cong x.\frac{n+1}{C} + o(x)$$

if x is small, where n is a geometric random variable.





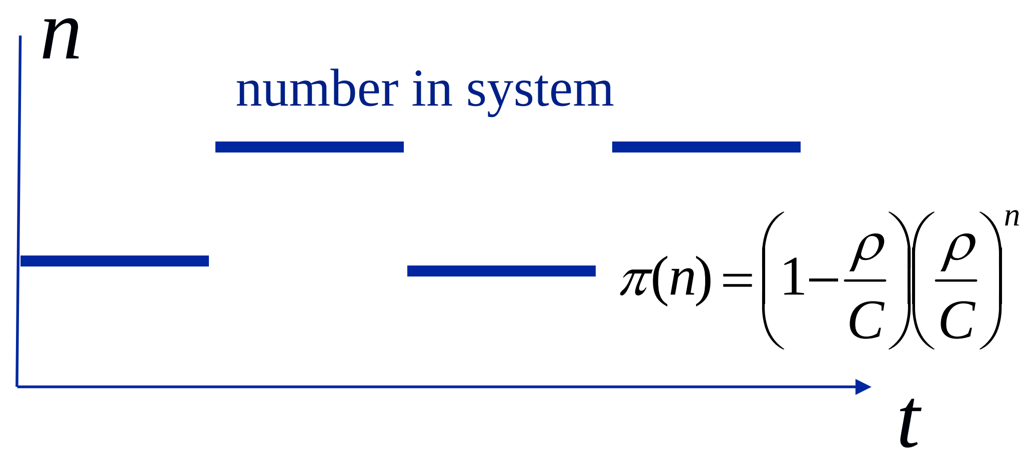
The M/G/1 processor sharing queue

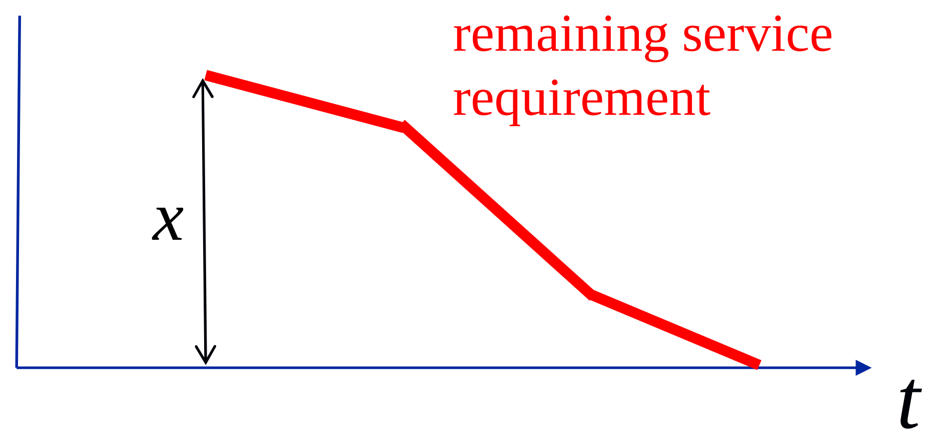
$$[S \mid x] \cong \frac{x}{C-\rho} + o(1/x)$$

if x is large;

$$[S \mid x] \cong x.\frac{n+1}{C} + o(x)$$

if x is small, where n is a geometric random variable.





$$E[S \mid x] = \frac{x}{C-\rho}$$

in both cases, of course!

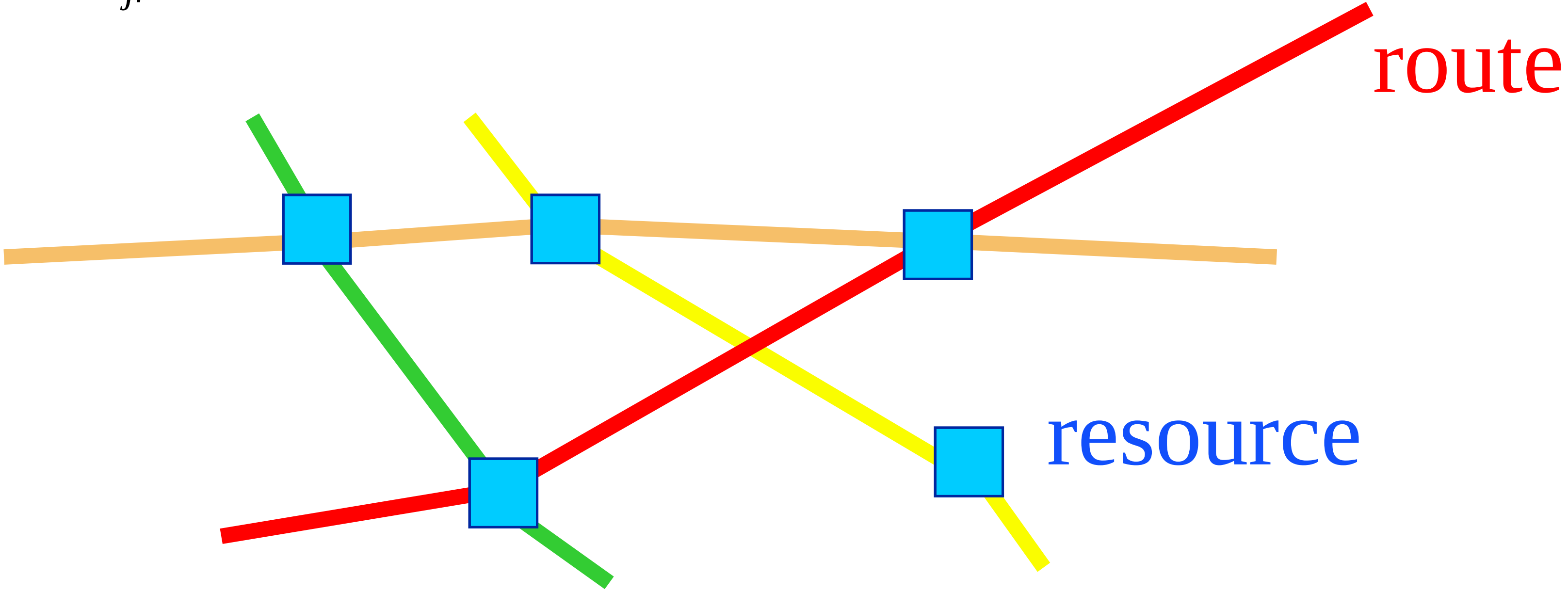
What is the network equivalent?

J - set of resources

R - set of routes

$A_{jr} = 1$ - if resource j is on route r

$A_{jr} = 0$ - otherwise



Rate allocation

n_r - number of flows on route r

x_r - rate of each flow on route r

Given the vector how are the rates chosen ?

$$n = (n_r, r \in R)$$

$$x = (x_r, r \in R)$$

Optimization formulation

Suppose $x = x(n)$ is chosen to

maximize $$\sum_r w_r n_r \frac{x_r^{1-\alpha}}{1-\alpha}$$

subject to $$\sum_r A_{jr} n_r x_r \le C_j \quad j \in J$$

$$x_r \ge 0 \quad r \in R$$

(weighted α-fair allocations, Mo and Walrand 2000)

$0 < \alpha < \infty$ (replace $\dfrac{x_r^{1-\alpha}}{1-\alpha}$ by $\log(x_r)$ if $\alpha = 1$)

Solution

$$x_r = \left(\frac{w_r}{\sum_j A_{jr} p_j(n)} \right)^{1/\alpha} \qquad r \in R$$

where

$$\sum_r A_{jr} n_r x_r \le \quad C_j \quad j \in J; \quad x_r \ge 0 \quad r \in R$$

$$p_j(n) \ge 0 \quad j \in J$$

$$p_j(n) \left(C_j - \sum_r A_{jr} n_r x_r \right) \ge 0 \quad j \in J$$

KKT conditions

$p_j(n)$ - *shadow price* (Lagrange multiplier) for the resource j capacity constraint

Examples of α-fair allocations

maximize $\displaystyle \sum_r w_r n_r \frac{x_r^{1-\alpha}}{1-\alpha}$

subject to $\displaystyle \sum_r A_{jr} n_r x_r \leq C_j \quad j \in J$

$x_r \geq 0 \quad r \in R$

$$x_r = \left(\frac{w_r}{\sum_j A_{jr} p_j(n)} \right)^{1/\alpha} \quad r \in R$$

$\alpha \to 0 \quad (w = 1)$ - maximum flow

$\alpha \to 1 \quad (w = 1)$ - proportionally fair

$\alpha = 2 \quad (w_r = 1/T_r^2)$ - TCP fair

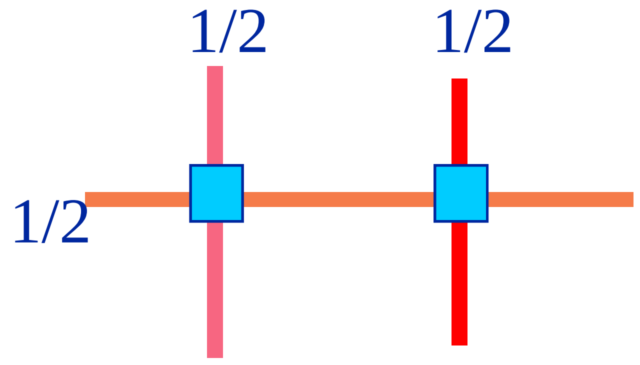
$\alpha \to \infty \quad (w = 1)$ - max-min fair

Example

$$n_r = 1,\ \ w_r = 1\ \ r \in R,$$
$$C_j = 1\ \ j \in J$$

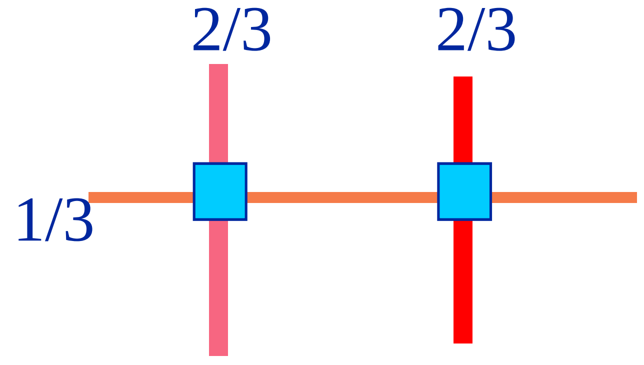
max-min fairness:

$$\alpha \to \infty$$



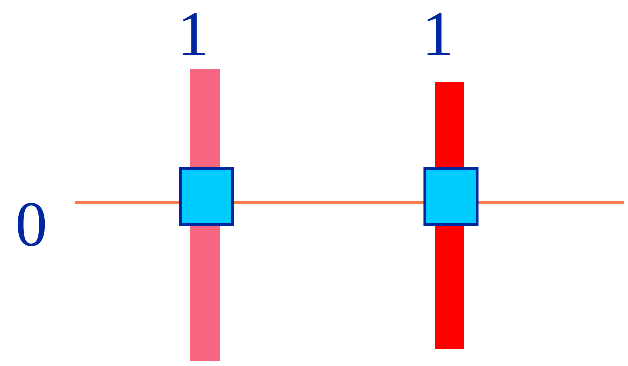
proportional fairness:

$$\alpha = 1$$



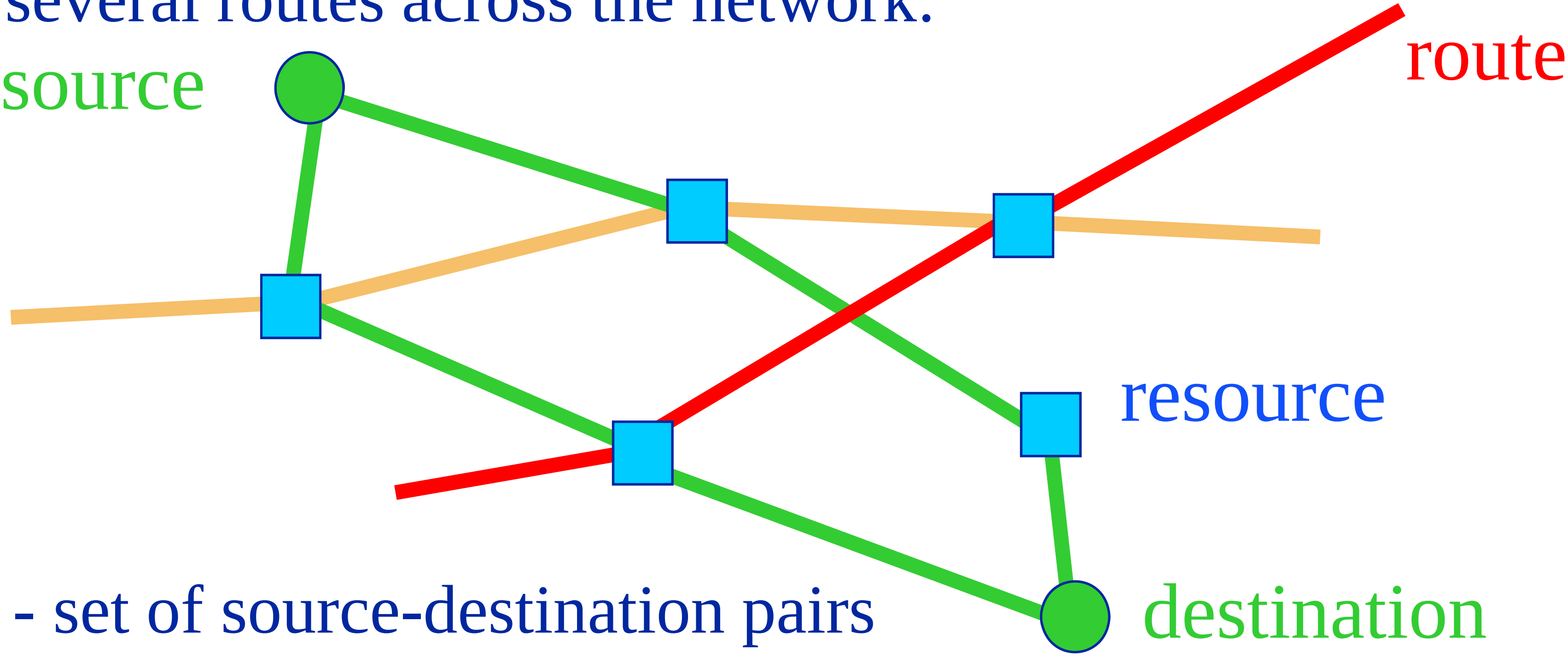
maximum flow:

$$\alpha \to 0$$



Multipath routing

Suppose a source-destination pair has access to several routes across the network:



S - set of source-destination pairs

$r \in s$ - route r serves s-d pair s

Combined multipath routing and congestion control: a robust Internet architecture. Key, Massoulié & Towsley

Routing and optimization formulation

Suppose $x = x(n)$ is chosen to

maximize $$\sum_s n_s \log(x_s)$$

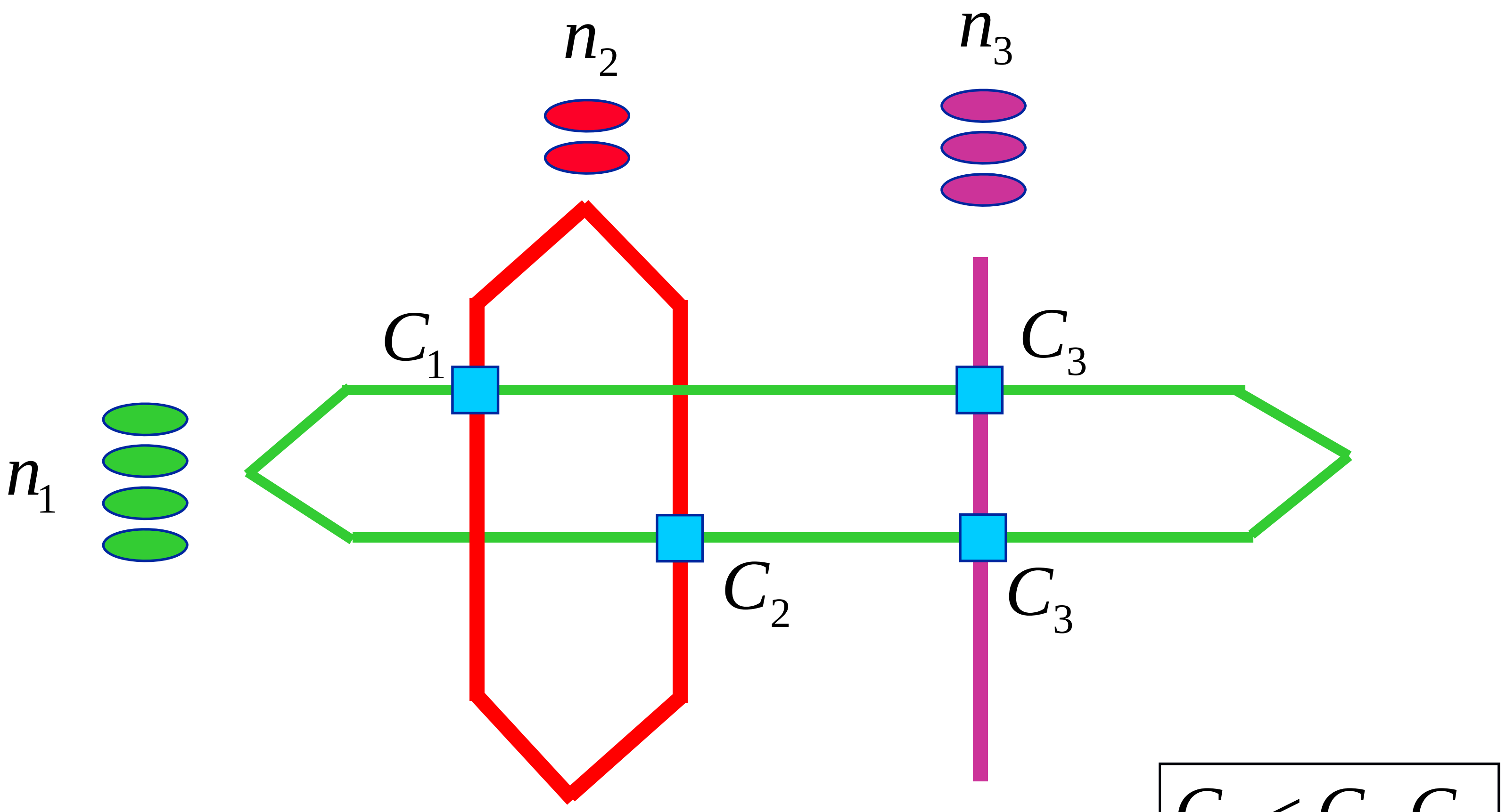
subject to $$\sum_r H_{sr}\, y_r = x_s \quad s \in S$$

$$\sum_r A_{jr}\, n_r\, y_r \le C_j \quad j \in J$$

$$y_r \ge 0 \quad r \in R$$

(H is an incidence matrix, showing which routes serve a source-destination pair)

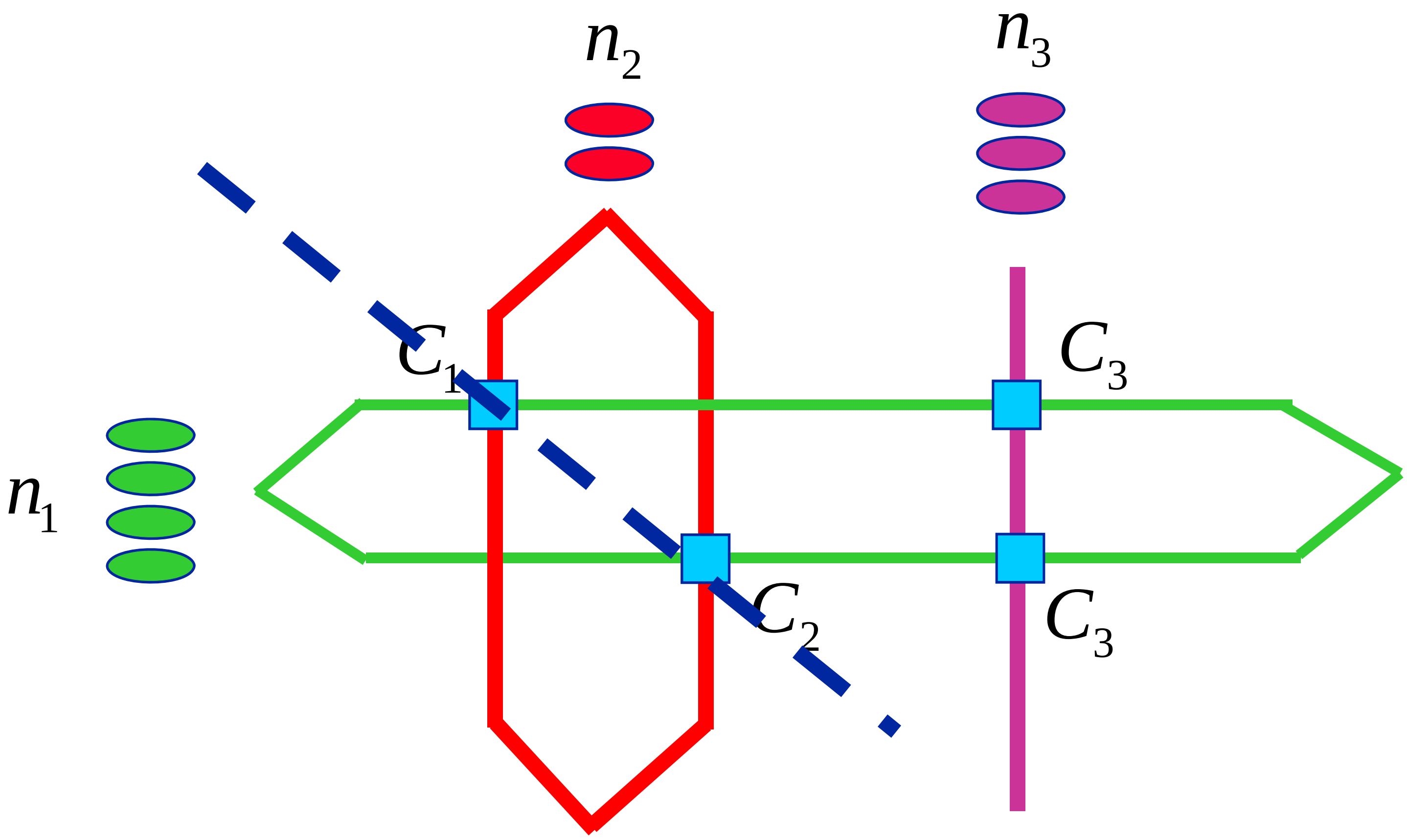
Example of multipath routing



Routes, as well as flow rates, are chosen to optimize

$\sum_s n_s \log(x_s)$ over source-destination pairs s

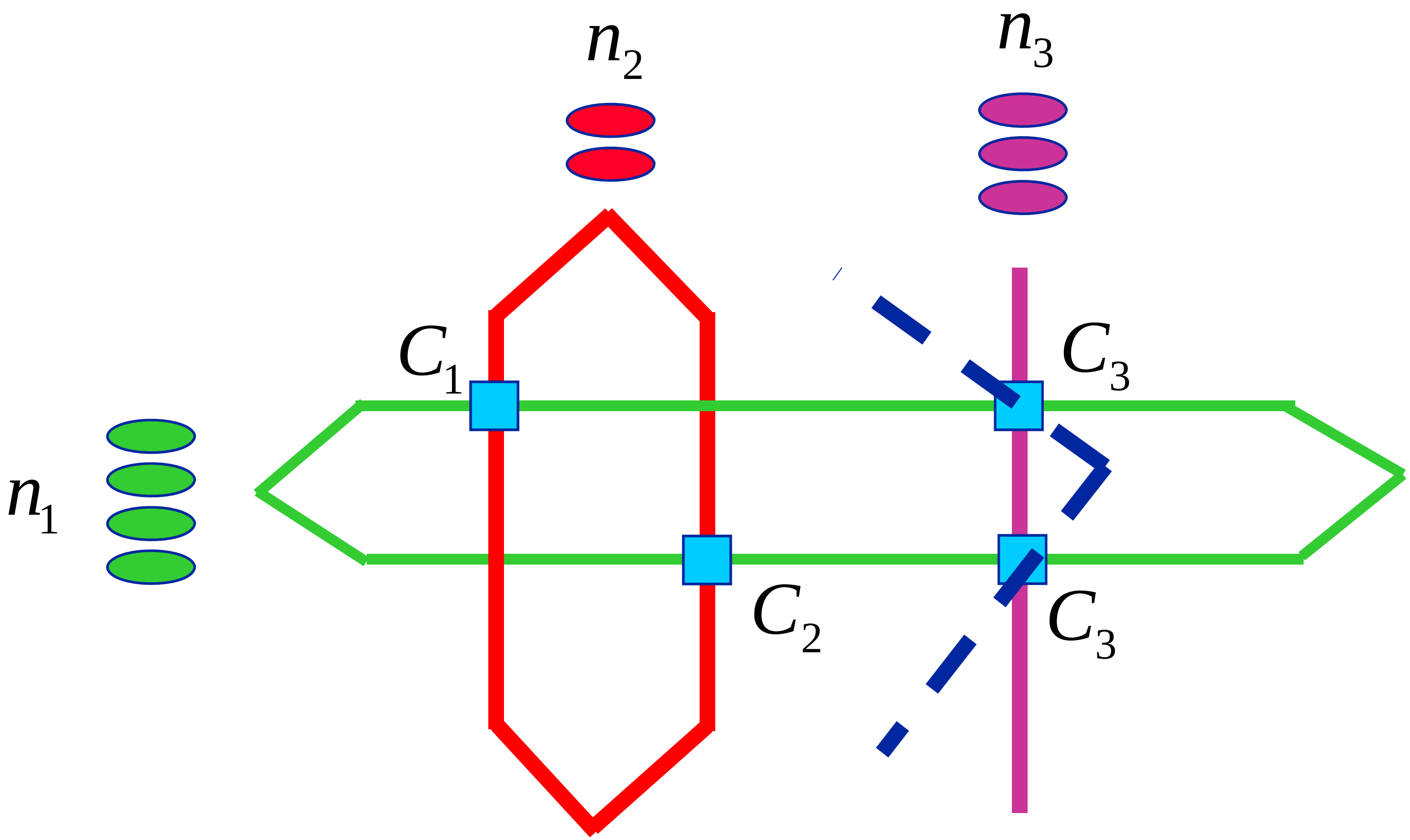
First cut constraint



$$n_1 x_1 + n_2 x_2 \leq C_1 + C_2$$

Cut defines a single *pooled resource*

Second cut constraint



$$\frac{1}{2}n_1x_1 + n_3x_3 \leq C_3$$

Cut defines a *second* pooled resource

Routing and optimization formulation

We may suppose $x = x(n)$ is chosen to

maximize $$\sum_s n_s \log(x_s)$$

subject to $$\sum_s \bar{A}_{js}\, n_s\, x_s \le \bar{C}_j \quad j \in \bar{J}$$

$$x_s \ge 0 \quad s \in S$$

where $\bar{J}$ is the set of pooled resources,
and $\bar{A}$ has non-negative entries

Proportional fairness

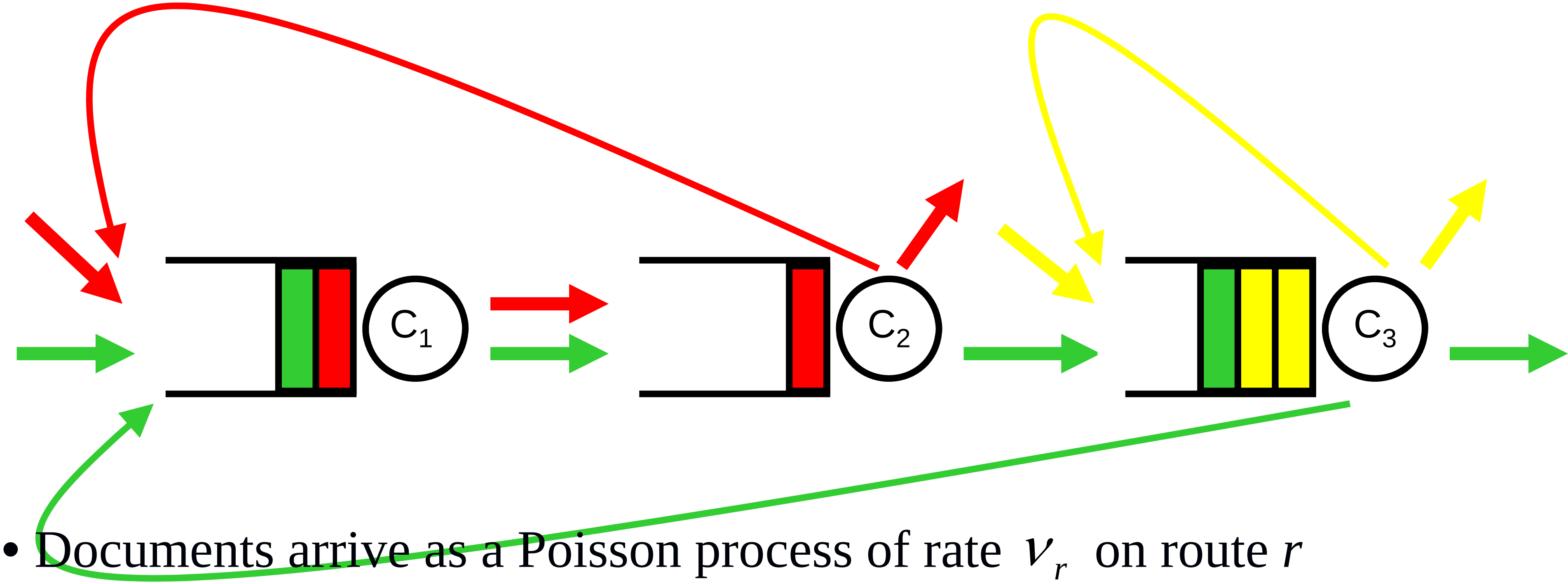
Henceforth we specialize to the case of proportional fairness, $\alpha = 1$, $w = 1$.

This case has interpretations in terms of axiomatic definitions of fairness, bargaining games, and distributed pricing.

Our aim is to explore the stochastic flow level model, to see if it shares some of the features of single resource processor sharing.

Why might one think it might?

A queueing network



• Documents arrive as a Poisson process of rate ν_r on route r
• Documents comprise an arbitrarily distributed number of packets
• These packets are transferred one by one through the network
• Packets have an arbitrary phase-type distribution of service requirement, which can differ from queue to queue
• each queue has a processor sharing discipline

Massoulié, Proutière 2003, Bonald and Proutière 2004,Walton 2009

Flow level model

Define a Markov process $n(t) = (n_r(t), r \in R)$

with transition rates

$n_r \to n_r + 1$ at rate ν_r $r \in R$

$n_r \to n_r - 1$ at rate $n_r \, x_r(n) \, \mu_r$ $r \in R$

- Poisson arrivals, exponentially distributed file sizes
- model originally due to Roberts and Massoulié 1998

Stability

Let $$\rho_r = \frac{\nu_r}{\mu_r} \quad r \in R$$

If $$\sum_r A_{jr} \rho_r < C_j \quad j \in J$$

then the Markov chain $n(t) = (n_r(t), r \in R)$
is positive recurrent

De Veciana, Lee & Konstantopoulos 1999;
Bonald & Massoulié 2001

Heavy traffic

We're interested in what happens when we approach the edge of the achievable region, when

$$\sum_r A_{jr}\,\rho_r \;\approx\; C_j \qquad j \in J$$

Fluid model for a network operating under a fair bandwidth-sharing policy. K & Williams *Ann Appl Prob 2004*

Product form stationary distributions for diffusion approximations to a flow level model operating under a proportional fair sharing policy. Kang, K, Lee & Williams *Performance Evaluation Review 2007*

State space collapse and diffusion approximation for a network operating under a proportional fair sharing policy. Kang, K, Lee & Williams *Ann Appl Prob* to appear

Fluid and diffusion scalings

Consider a sequence of networks, labelled by N, where as $N \to \infty$,

$$\nu^N \to \nu, \quad \mu^N \to \mu, \quad N(A\rho^N - C) \to \theta$$

(and thus $A\rho = C$)

Fluid scaling:

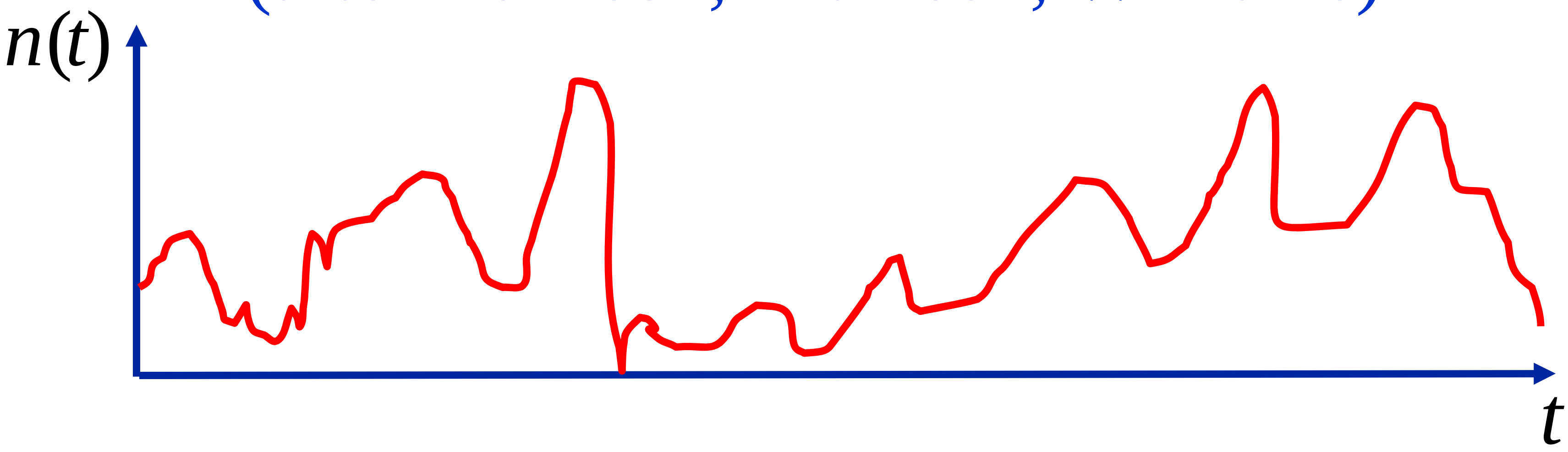
$$\frac{n^N(Nt)}{N}$$

Diffusion scaling:

$$\frac{n^N(N^2 t)}{N}$$

Fluid and diffusion scalings

(after Harrison, Bramson, Williams)



Fluid scaling:

$$\frac{n^N(Nt)}{N}$$

On this time scale, traffic and capacity are balanced, and we expect a law of large numbers

Diffusion scaling:

$$\frac{n^N(N^2t)}{N}$$

On this time scale, there is a drift of θ, and we expect a central limit theorem

Balanced fluid model

Suppose $\sum_r A_{jr}\ \rho_r \quad = \quad C_j \qquad j \in J$

and consider differential equations

$$\frac{\mathrm{d}n_r(t)}{\mathrm{d}t} = \nu_r - n_r x_r(n)\mu_r \qquad (n_r > 0) \qquad r \in R$$

First let's substitute for the values of $\quad x_r(n), \quad r \in R$, to give:

$$\frac{\mathrm{d}n_r(t)}{\mathrm{d}t} = \nu_r - \frac{n_r \mu_r}{\sum_j A_{jr} p_j(n)} \qquad r \in R$$

(care needed when $n_r = 0$).

Thus, at an invariant state,

$$n_r = \frac{\nu_r}{\mu_r} \sum_j A_{jr} p_j(n) \qquad r \in R$$

State space collapse: invariant manifold

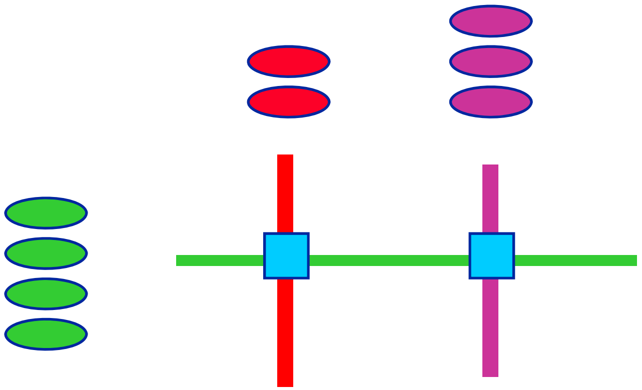
The following are equivalent:

- n is an invariant state
- there exists a non-negative vector p with

$$n_r = \frac{\nu_r}{\mu_r} \sum_j A_{jr} p_j \quad r \in R$$

Thus the set of invariant states forms a J dimensional subspace, parameterized by p.

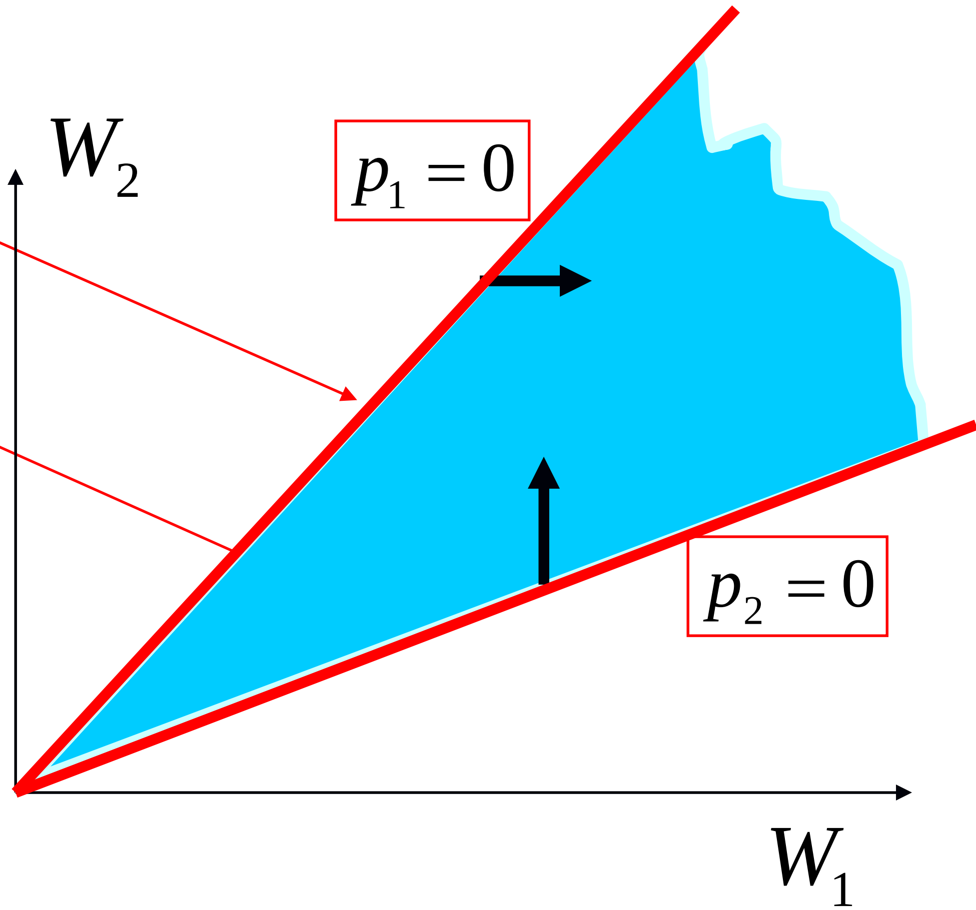
Example



$$\mu_r = 1, \quad r \in R$$

slope $\dfrac{\rho_2 + \rho_0}{\rho_0}$

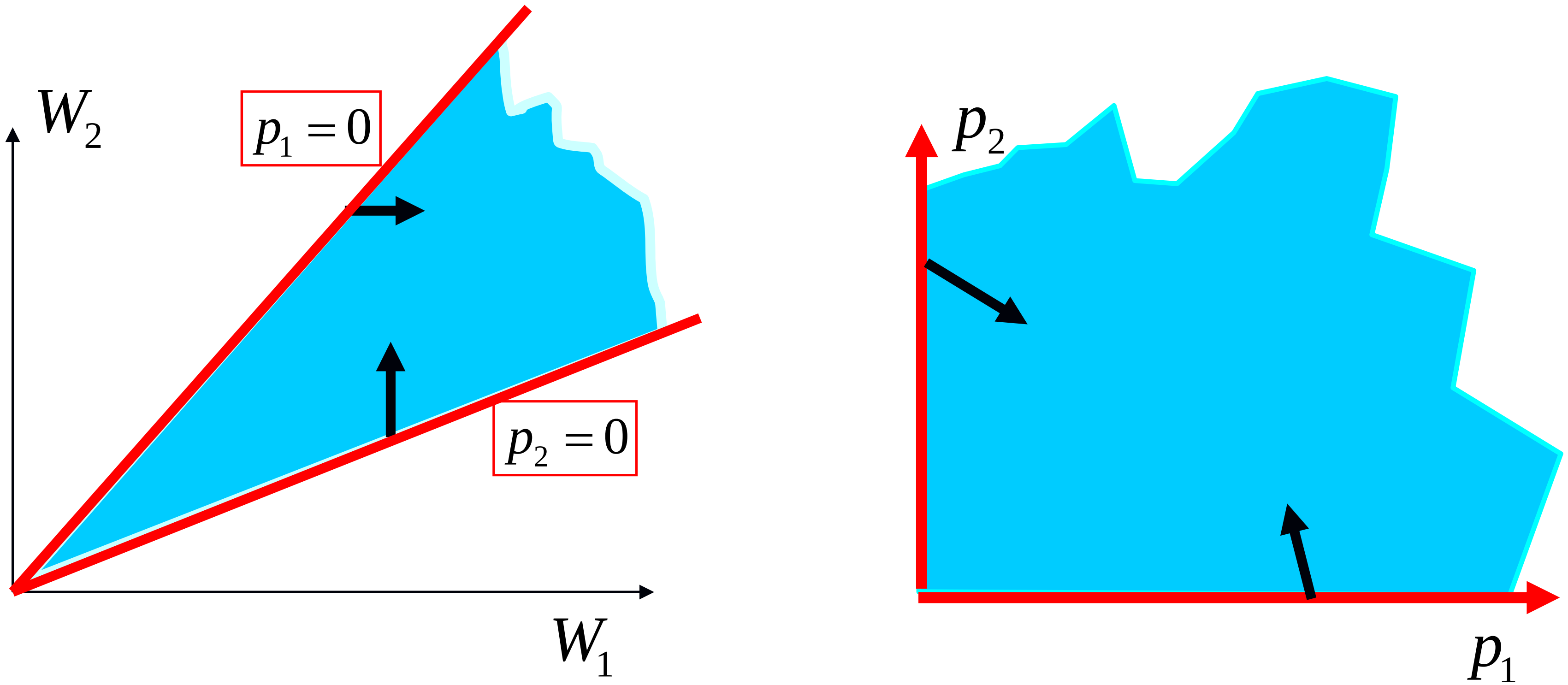
slope $\dfrac{\rho_0}{\rho_1 + \rho_0}$



Each bounding face corresponds to a resource not working at full capacity

Entrainment: congestion at some resources may prevent other resources from working at their full capacity.

Stationary distribution?



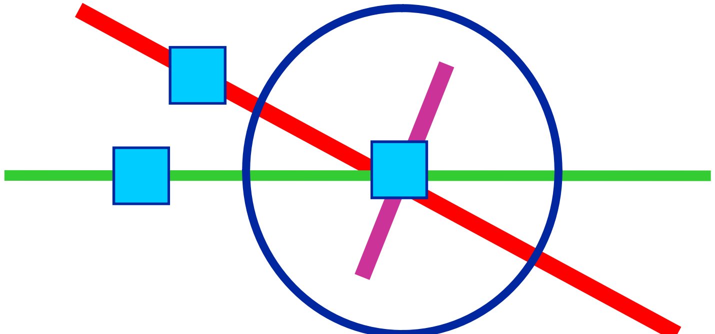
Williams (1987) determined sufficient conditions, in terms of the reflection angles and covariance matrix, for a SRBM in a polyhedral domain to have a product form invariant distribution – a skew symmetry condition

Local traffic condition

Assume the matrix A contains the columns of the unit matrix amongst its columns:

$$A = \begin{pmatrix} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & 0 & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & 1 & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & 0 & 1 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & 0 & 0 & 1 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

i.e. each resource has some local traffic -



Product form under proportional fairness

$$\alpha = 1, w_r = 1, \; r \in R$$

Under the stationary distribution for the reflected Brownian motion, the (scaled) components of p are independent and exponentially distributed. The corresponding approximation for n is

$$n_r \approx \rho_r \sum_j A_{jr} p_j \quad r \in R$$

where

$$p_j \sim \mathrm{Exp}(C_j - \sum_r A_{jr} \rho_r) \quad j \in J$$

Dual random variables are independent and exponential

Product form under proportional fairness

In general, stability requires

$$\sum_s \overline{A}_{js}\ \rho_s \quad < \quad \overline{C}_j \qquad j \in \overline{J}$$

- a collection of *generalized cut constraints*.

Provided $\overline{A}$ contains a unit matrix, we have the approximation

$$n_s \approx \rho_s \sum_{j \in \overline{J}} \overline{A}_{js} p_j \qquad s \in S$$

where

$$p_j \sim \text{Exp}(\overline{C}_j - \sum_s \overline{A}_{js} \rho_s) \qquad j \in \overline{J}$$

Independent dual random variables, one for each generalized cut constraint – network generalization of processor sharing

Processor sharing for a network?

Large deviations and heavy traffic point to subtly different results in the case where the matrix A does not contain a unit matrix – this case is important for resource pooling applications.

Challenge: establish straightforwardly the approximation -

$$E\left[\text{sojourn time}, S, \text{on route } r \;\middle|\; \text{job size}, x\right]$$

$$= x \sum_j \frac{A_{jr}}{C_j - \sum_{r'} A_{jr'} \rho_{r'}}$$