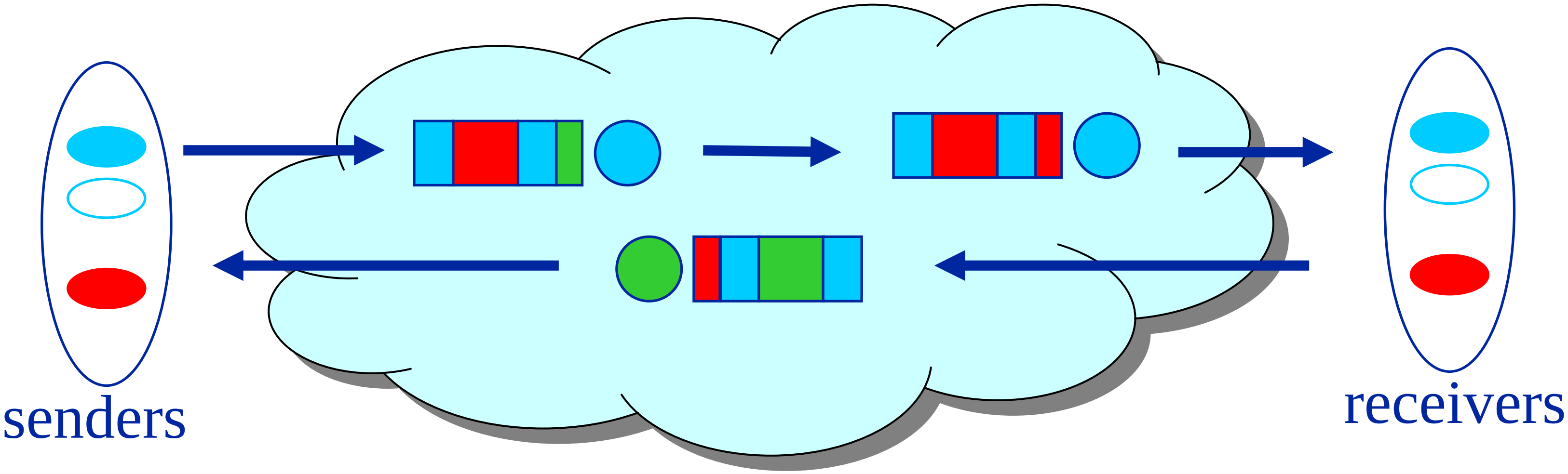
Models of network routing and congestion control

Frank Kelly, Cambridge

www.statslab.cam.ac.uk/~frank/TALKS/amherst.html

University of Massachusetts Amherst,
March 26, 2008

End-to-end congestion control



Senders learn (through feedback from receivers) of congestion at queue, and slow down or speed up accordingly. With current TCP, throughput of a flow is proportional to

$$1/(T\sqrt{p})$$

T = round-trip time, p = packet drop probability.

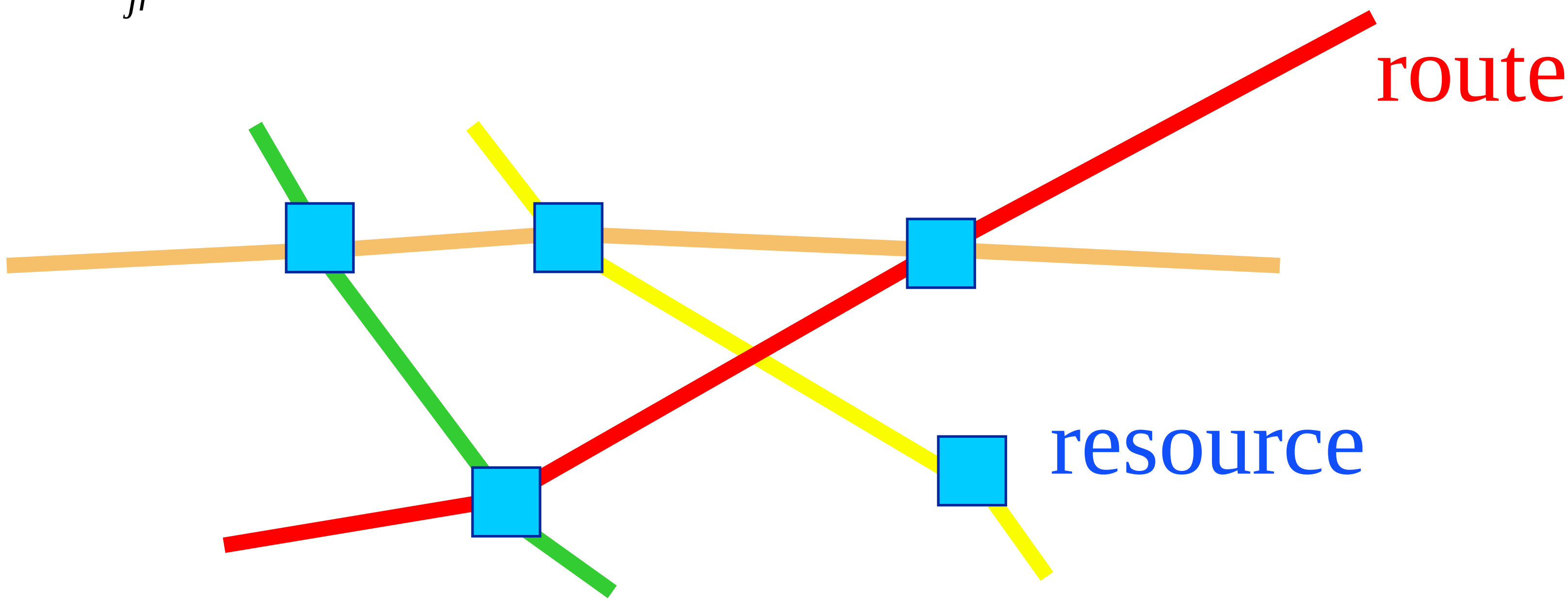
(Jacobson 1988, Mathis, Semke, Mahdavi, Ott 1997, Padhye, Firoiu, Towsley, Kurose 1998, Floyd & Fall 1999)

Model definition

- We want to describe a network model, with fluctuating numbers of flows
- We first need
  - notation for network structure
  - abstraction of rate allocation
- Then we need to define the random nature of flow arrivals and departures

Network structure (J, R, A)

J	-	set of resources
R	-	set of routes
$A_{jr} = 1$	-	if resource j is on route r
$A_{jr} = 0$	-	otherwise



Rate allocation

w_r - weight of route r

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 $\frac{x_r^{1-\alpha}}{1-\alpha}$ by $\log(x_r)$ if $\alpha = 1$)

Solution

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

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

Observe alignment with square-root formula when

$$\alpha = 2, \quad w_r = 1/T_r^2, \quad p_r \approx \sum_j A_{jr} p_j$$

Examples of α-fair allocations

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

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

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

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

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

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

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

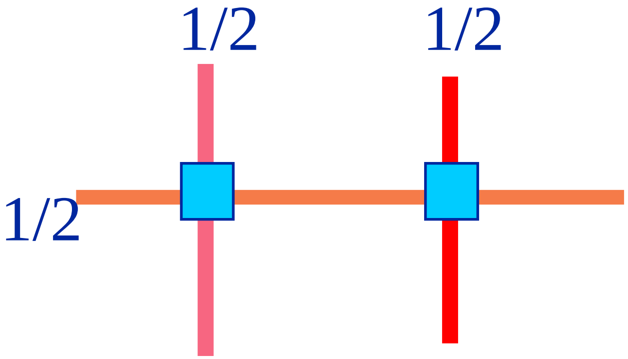
$\alpha \to \infty \quad (w = 1)$ - max-min fair

Example

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

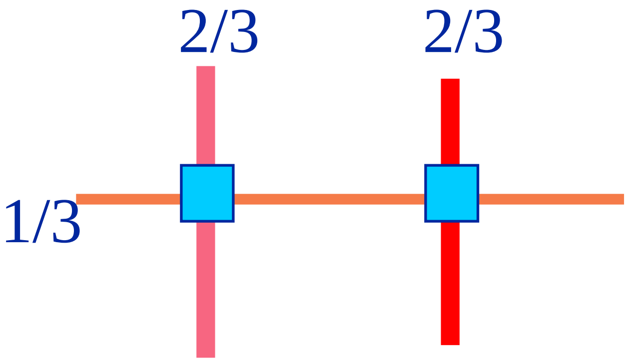
max-min fairness:

$$\alpha \to \infty$$



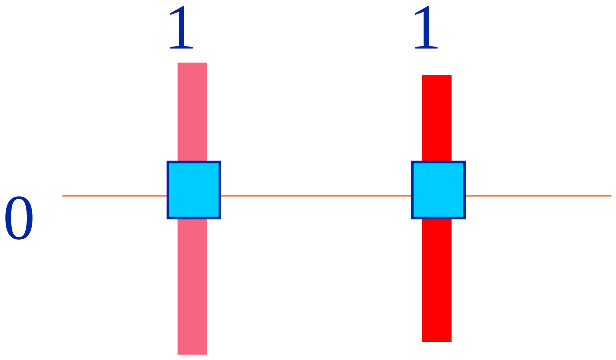
proportional fairness:

$$\alpha = 1$$



maximum flow:

$$\alpha \to 0$$



Flow level model

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

with transition rates

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

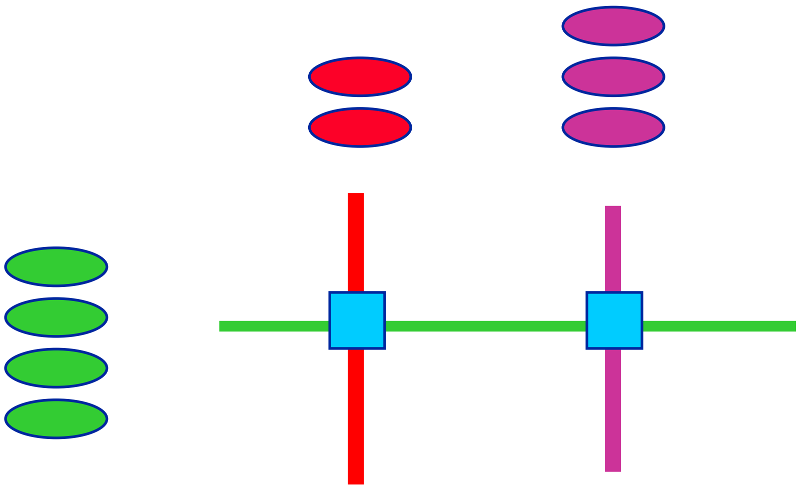
$n_r \to n_r - 1$ at rate $n_r \; x_r(n) \; \mu_r$ $r \in R$

- Poisson arrivals, exponentially distributed file sizes
- model originally due to Roberts and Massoulié 1998
- for a single resource (or a linear network with proportional fairness) we can allow arbitrary file size distributions – becomes a quasi-reversible node

Example: a linear network

$$\alpha = 1, \ C_j = 1 \quad j \in J$$

$$w_r = 1, \quad \rho_r = \nu_r / \mu_r \quad r \in R$$

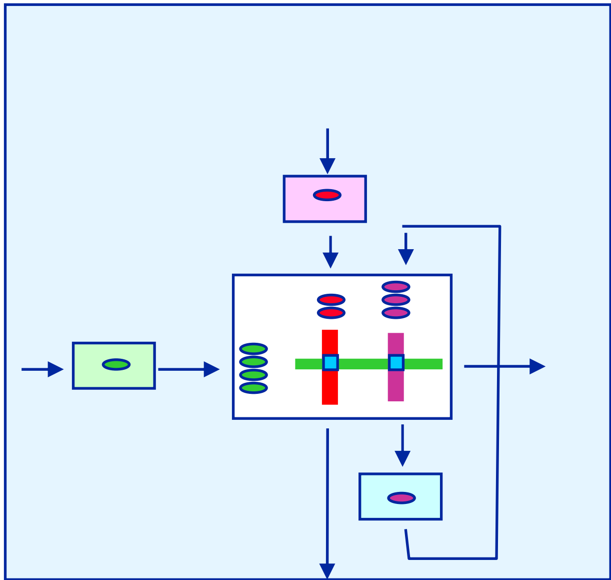


Quasi-reversible,

with:

$$\pi(n_0, n_1, n_2) = B \binom{\sum_0^2 n_r}{n_0} \prod_0^2 \rho_r^{n_r}$$

$$B = (1-\rho_0)^{-1} \prod_1^2 (1-\rho_0-\rho_r)$$



Stability

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

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

and resource allocation is weighted α-fair

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

is positive recurrent

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

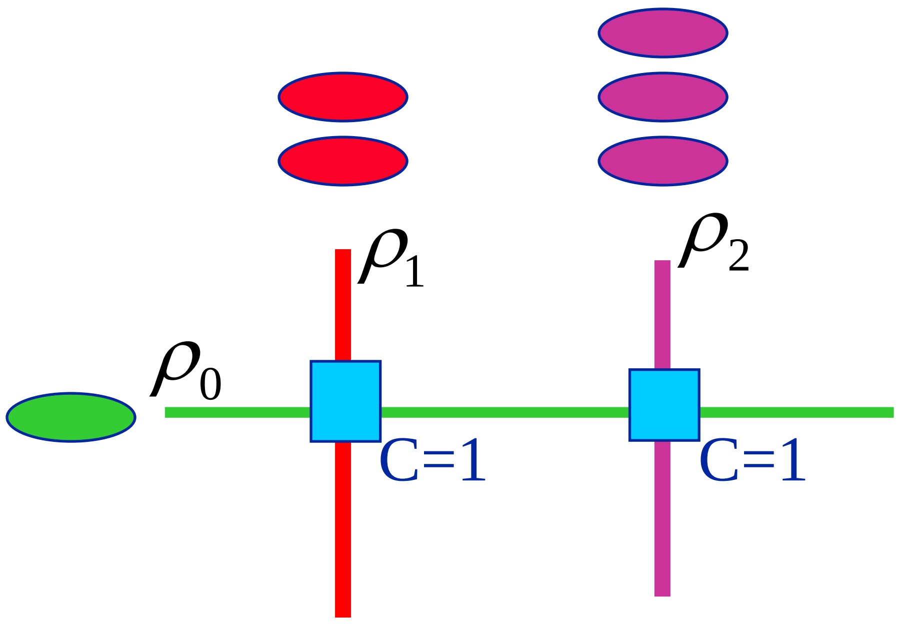
What goes wrong without fairness?

Suppose vertical streams have priority: then condition for stability is

$$\rho_0 < (1-\rho_1)\ (1-\rho_2)$$

and *not*

$$\rho_0 < \min\{1-\rho_1,\ 1-\rho_2\}$$



Bonald & Massoulié 2001

Heavy traffic

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

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

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

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 $A\rho = C$)

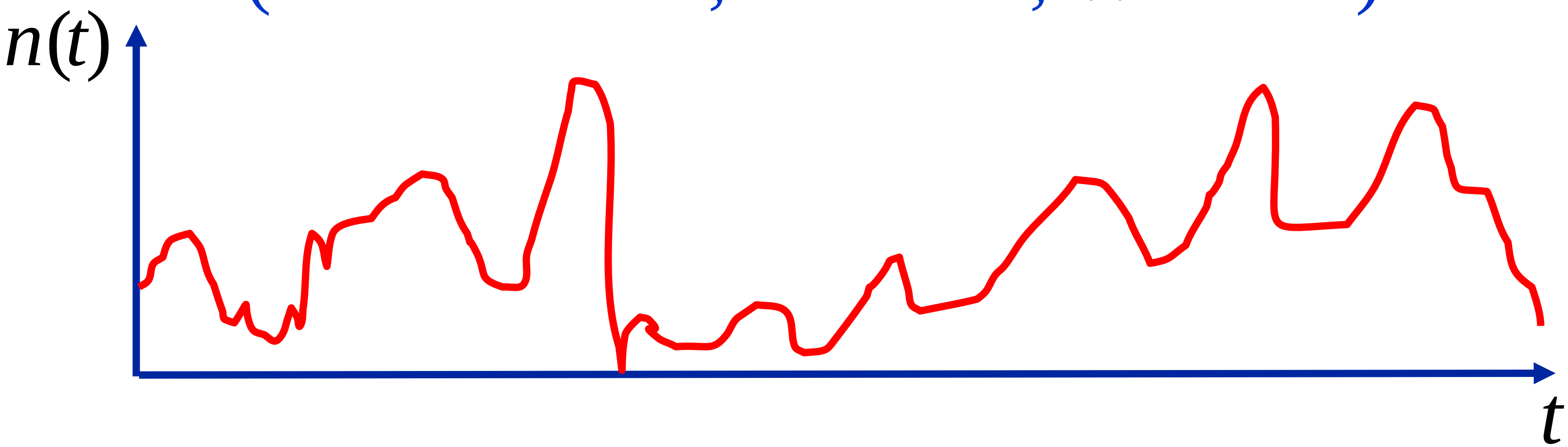
Fluid scaling:

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

Diffusion scaling:

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

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:

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

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 $x_r(n), \quad r \in R$, to give:

$$\frac{\mathrm{d}n_r(t)}{\mathrm{d}t} = \nu_r - n_r\mu_r \left(\frac{w_r}{\sum_j A_{jr} p_j(n)} \right)^{1/\alpha} \qquad r \in R$$

(care needed when $n_r = 0$).

Thus, at an invariant state,

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

State space collapse: invariant manifold

The following are equivalent:

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

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

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

A potential function

Let $$F(n) = \frac{1}{\alpha+1} \sum_r \nu_r w_r \mu_r^{\alpha-1} \left(\frac{n_r}{\nu_r} \right)^{\alpha+1}$$

(following Bonald and Massoulié 2001). Then

$$\frac{\mathrm{d}}{\mathrm{d}t} F(n(t)) \leq 0$$

with equality only if n is an invariant state.

Workloads

Let

$$W_j(n(t)) = \sum_r A_{jr} \frac{n_r(t)}{\mu_r}$$

the *workload* for resource j. Then

$$\frac{\mathrm{d}}{\mathrm{d}t} W_j(n(t)) \geq 0, \qquad p_j(n(t)) \frac{\mathrm{d}}{\mathrm{d}t} W_j(n(t)) = 0$$

Extremal characterization of an invariant state

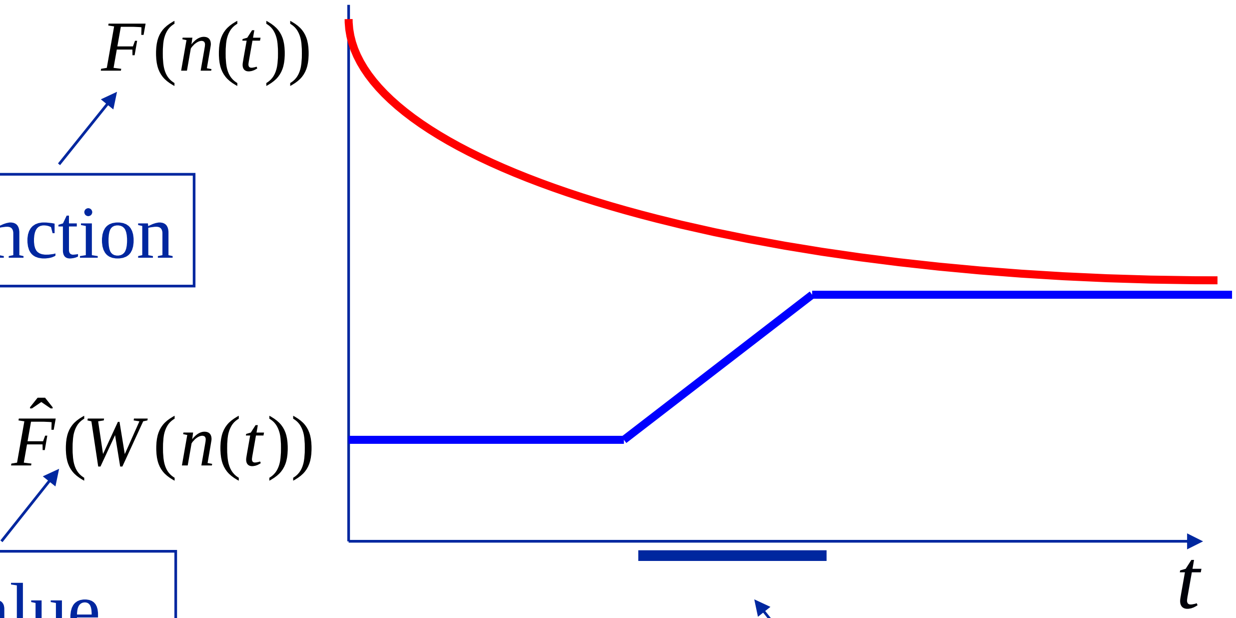
Minimize $$F(n) = \frac{1}{\alpha+1}\sum_r \nu_r w_r \mu_r^{\alpha-1}\left(\frac{n_r}{\nu_r}\right)^{\alpha+1}$$

subject to $$\sum_r A_{jr}\frac{n_r}{\mu_r} \ge W_j \quad j \in J, \quad n_r \ge 0 \quad r \in R$$

Solution is $$n_r = \frac{\nu_r}{\mu_r}\left(\frac{\sum_j A_{jr}\hat{p}_j(W)}{w_r}\right)^{1/\alpha} \qquad r \in R$$

$\hat{p}_j(W)$ - Lagrange multiplier for the resource j *workload* constraint

Evolution of functions *F*



potential function

extremal value, given workloads

period where workload is increasing

$$F(n(t)) - \hat{F}(W(n(t))$$

provides a Lyapunov function which shows convergence to the invariant manifold

The case $\alpha = 1$

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

Define diagonal matrices

$$\rho = diag(\nu_r / \mu_r, \ \ r \in R), w = diag(w_r, \ \ r \in R)$$

Then $\quad n = \rho w^{-1} A^T p$

and so $\quad W = (A\mu^{-1})n = (A\mu^{-1}\rho w^{-1} A^T) p,$

$$p = (A\mu^{-1}\rho w^{-1} A^T)^{-1} W$$

Thus W lies in the polyhedral cone

$$\{W: \;\; W = A\mu^{-1}\rho\, w^{-1}A^{T}p,\; p \geq 0\}$$

More generally, W lie in the cone

$$(A\mu^{-1}\rho\, w^{-1/\alpha})\,C_{\alpha}$$

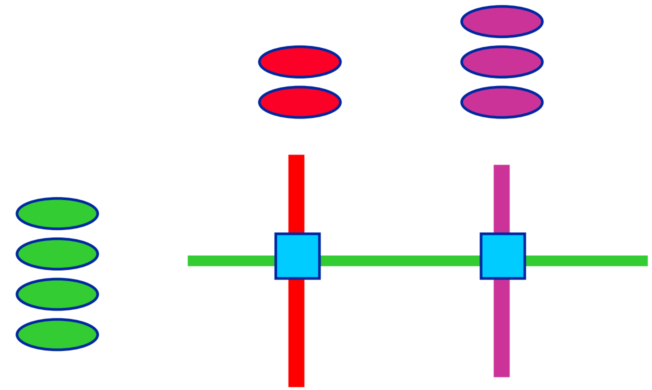
where

$$C_{\alpha} = \{(\sum_{r} A_{jr}\,p_{j})^{1/\alpha},\;\; r \in R\}$$

Example

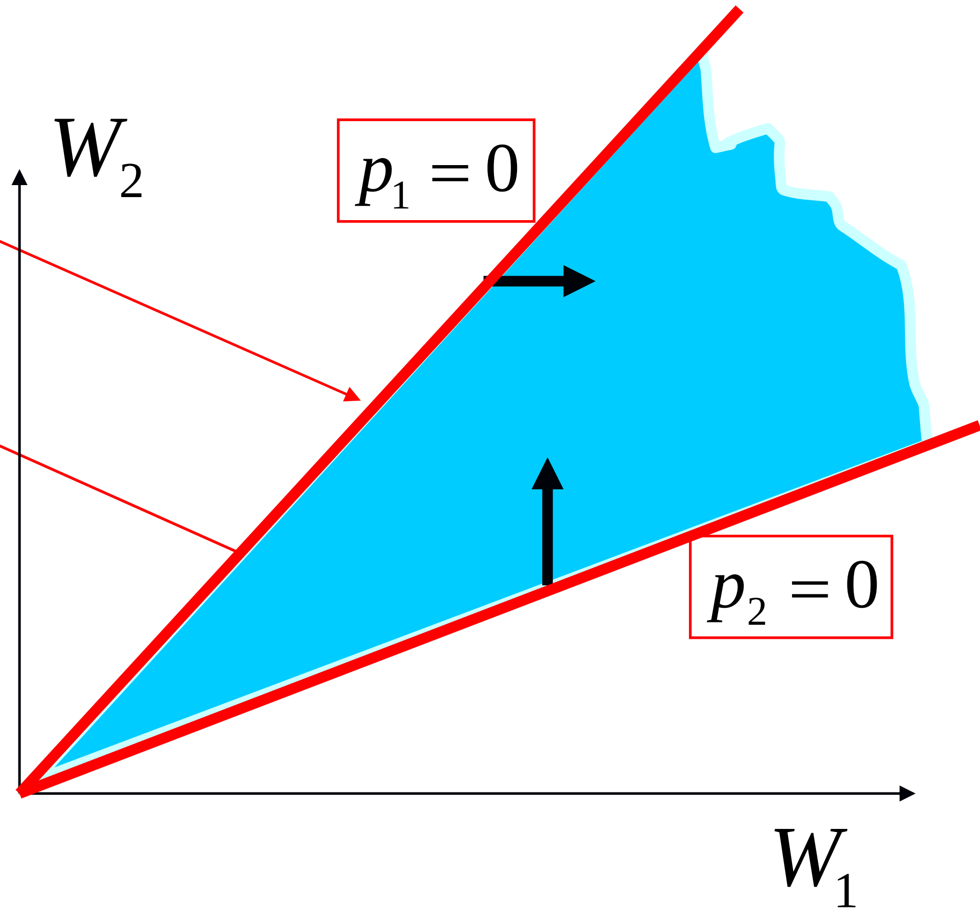
$$0 < \alpha < \infty$$

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



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

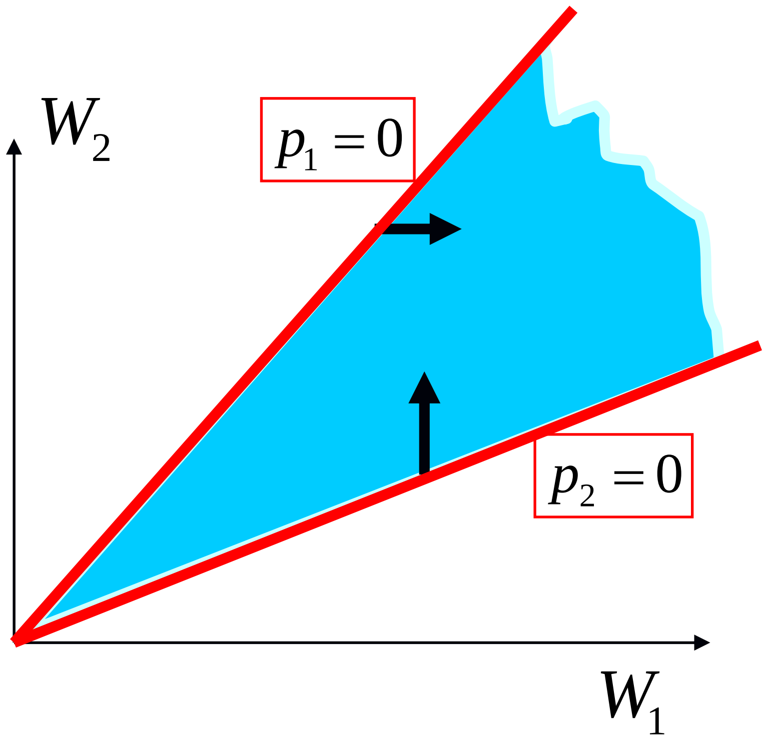
slope $\dfrac{\rho_0}{\rho_1 + \rho_0}$

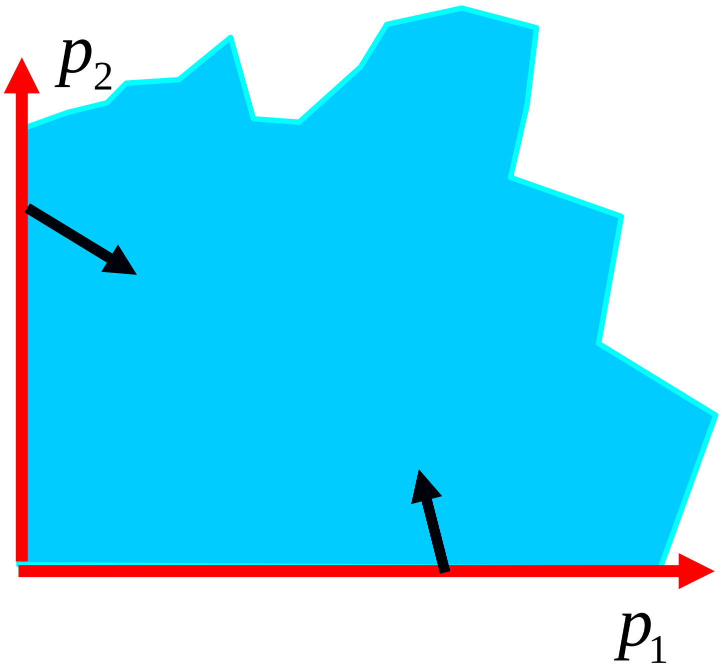


Each bounding face corresponds to a resource not working at full capacity

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

Stationary distribution?





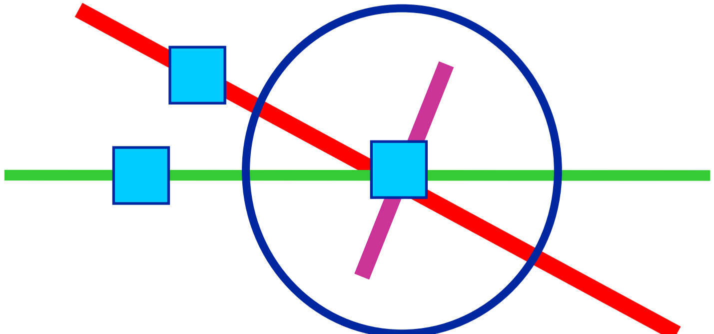
Williams (1987) determined sufficient conditions, in terms of the reflection angles and covariance matrix, for a SRBM in a polyhedral domain to have a product form invariant distribution – a skew symmetry condition

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

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

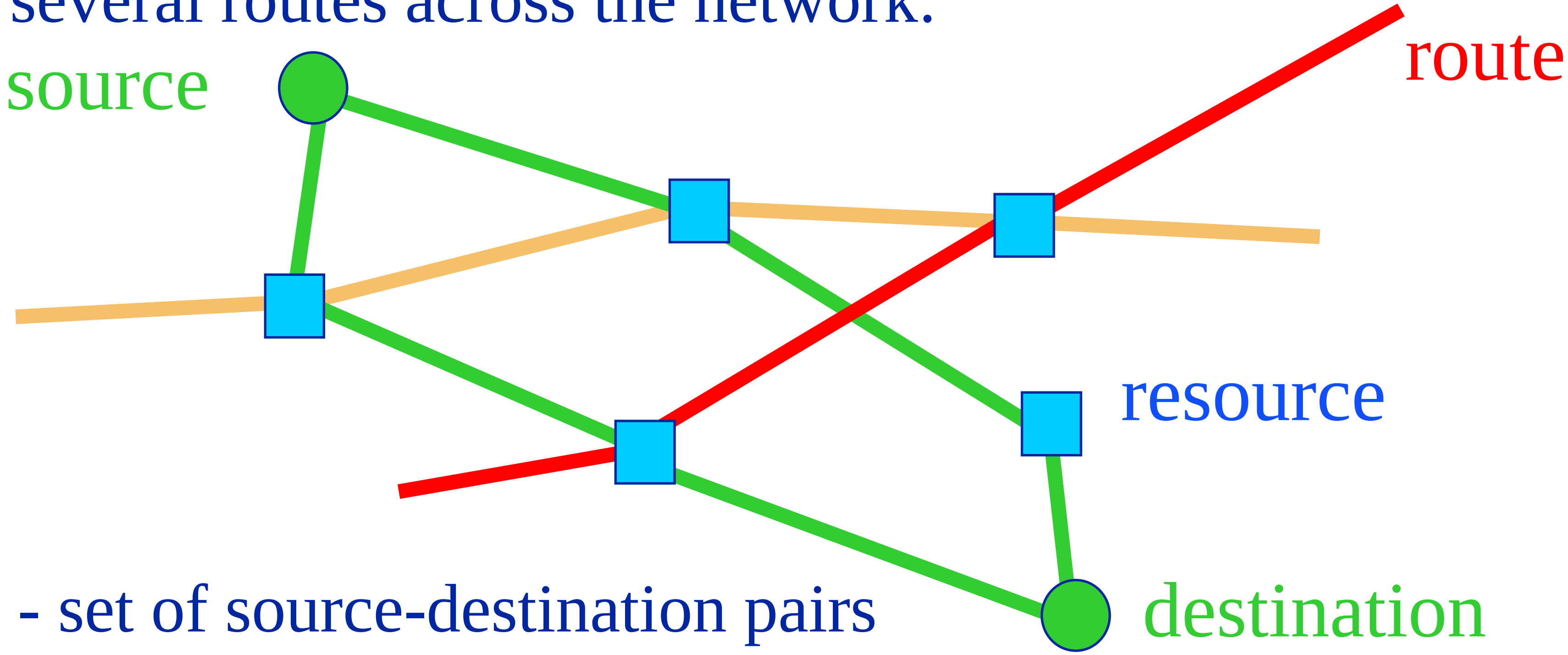
where

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

Dual random variables are independent and exponential

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

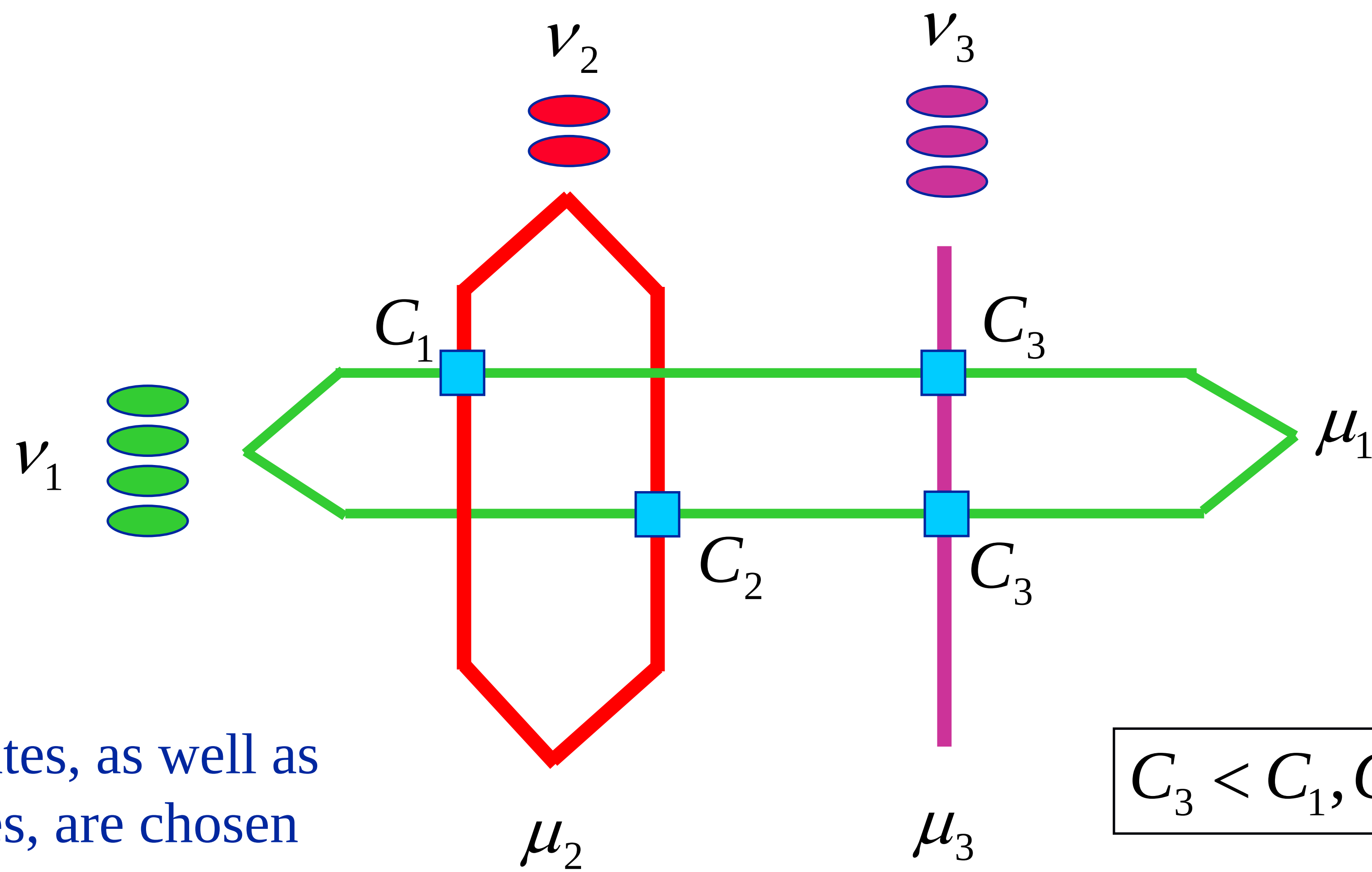
Suppose $x = x(n)$ is chosen to

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

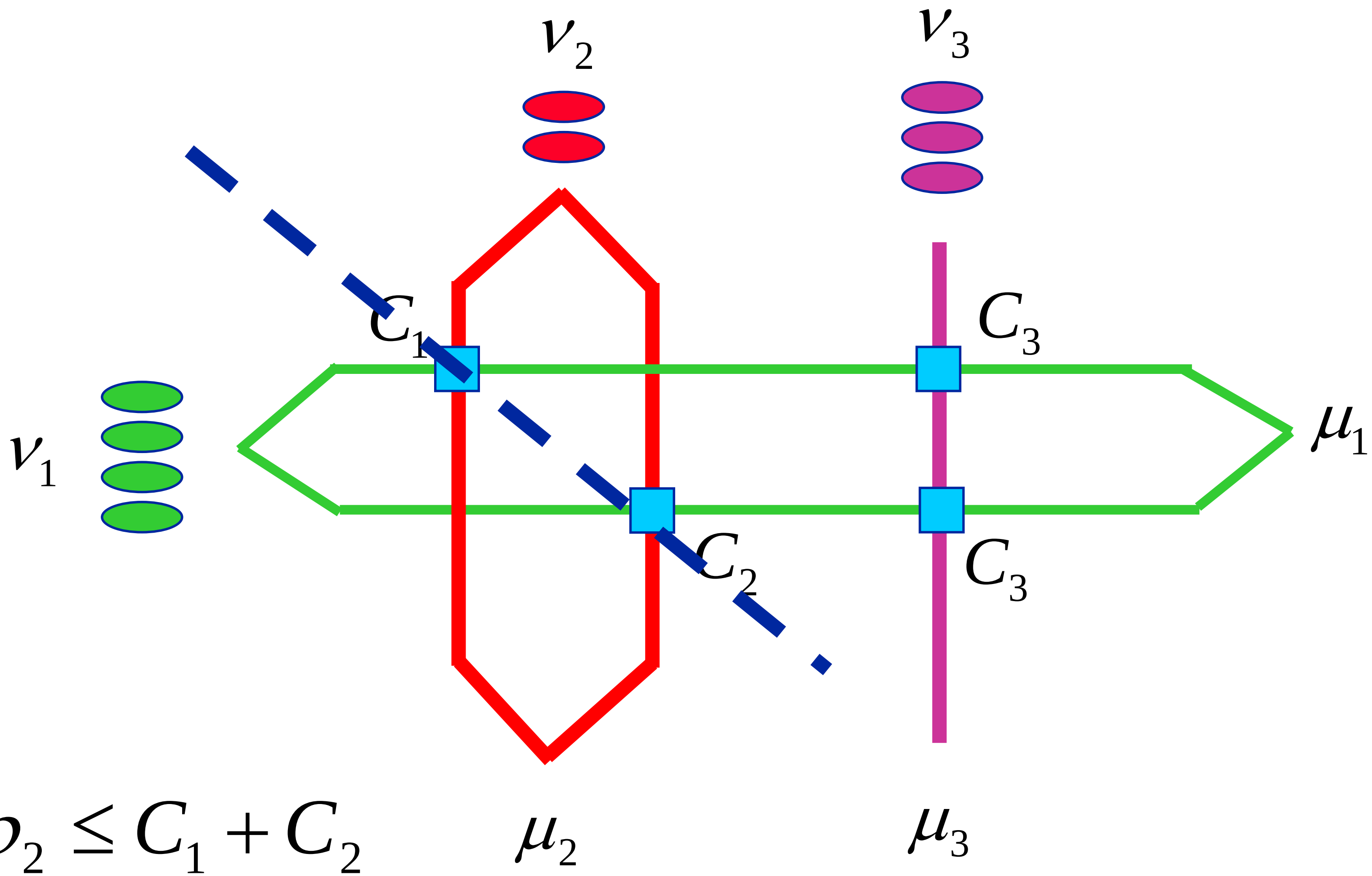


$C_3 < C_1, C_2$

Thus routes, as well as flow rates, are chosen to optimize

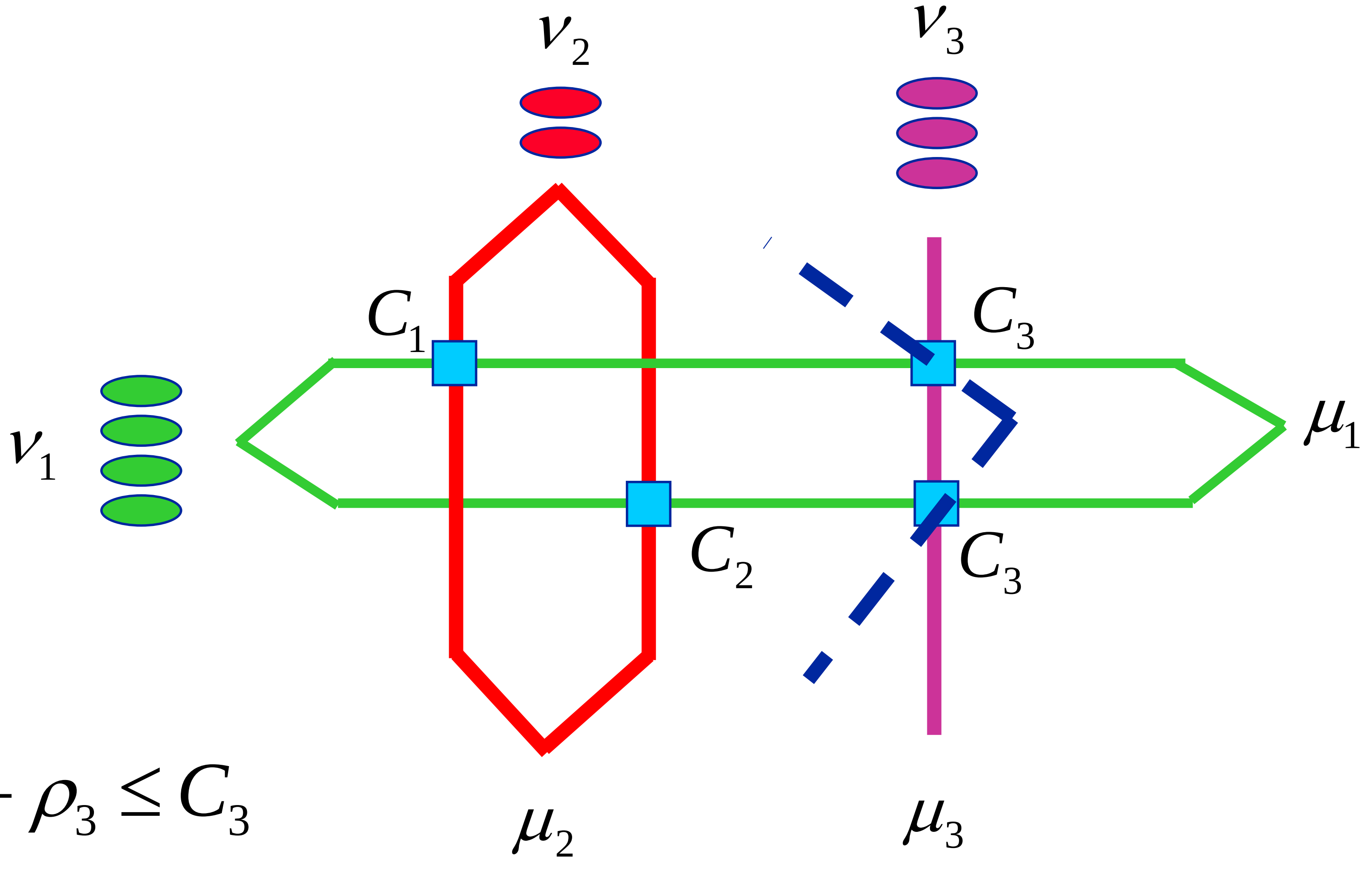
$\sum_s n_s \log(x_s)$ over source-sink pairs s

First cut constraint



$$\rho_1 + \rho_2 \leq C_1 + C_2$$

Second cut constraint



$$\frac{1}{2}\rho_1 + \rho_3 \leq C_3$$

Generalized cut constraints

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 again have the approximation

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

where

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

Again independent dual random variables, now one for each generalized cut constraint

Models of routing and congestion control

- Flow level Markov chain model
- Heavy traffic and proportional fairness give product form for dual variables
- A dual variable for each generalized cut constraint, under multipath routing
- Good behaviour, achieved without prior knowledge of which cut constraints bite