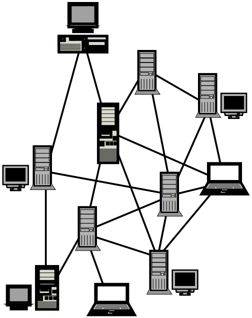
Load Balancing [AKU03, AKU05]



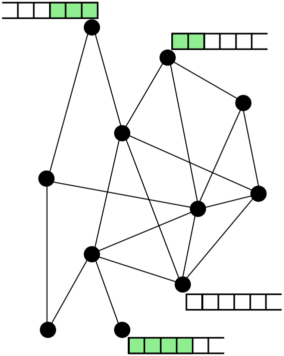
The problem:

- A network of n processors / servers
- Jobs are continuously generated on the processors
- The processors execute one job per step
- The processors can send any number of jobs to their neighbors
- There is no central control (decisions must be local)
- We study the long-term performance

Goals

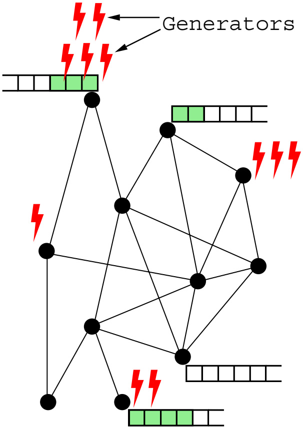
- **Stability**
  - We want the expected total load of the system to remain bounded w.r.t. time
  - Ideally polynomial in the system size
- **Efficiency (QoS)**
  - We want to minimize the waiting time in expectation and in probability

Network Model



- Arbitrary graph with n nodes
  - Nodes = Processors
  - Edges = Links
- Discrete time, $t = 0, 1, 2, \ldots$
- Every processor has a queue with jobs. In every step
  - It can send some jobs to a neighbor
  - It executes the first job in its queue
- Models networks where execution cost is significantly higher than transfer cost

Input Model



- In the beginning of every step, an adversary places *n* generators on the nodes
  - The adversary has full information of the history
  - Is not necessarily Markovian
- Every generator independently creates a new job with probability $\lambda < 1$
  - Expected number of new jobs $= \lambda n$

In Every Step:

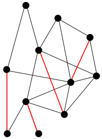
- The adversary places n generators on the nodes
- Every generator creates a job with probability $\lambda < 1$
- Load-balancing protocol
  - May move some jobs between neighbors
- Each nonempty node executes the first job in its queue

- If all nodes have jobs then
  - expected λn new jobs enter the system
  - n jobs leave

Protocol

In each step

- Choose a random matching in the graph
  - This can be done locally (in a distributed way).



- Equalize (modulo 1) the loads between the matched nodes
- Each node executes the first job in its queue

Analysis - a "Static" Invariant

- Split time into epochs of fixed length, $\tau = O(\log n)$
- An epoch is successful if, when it ends, the total load of the system is either
  - below a threshold $\Theta = O(n \log n)$, or
  - decreased by at least $\Delta = O(n \log n)$

Lemma

An epoch is successful with probability at least

$$1 - \frac{1}{n^c},$$

conditioned on all past events.

How to derive a dynamic result from the static invariant?

Drift Criteria

Definition

- A stochastic process $X_1, X_2, X_3 \ldots$ is a supermartingale if for all $n \geq 1$
$$\mathbf{E}[X_{n_1} \mid X_1, \ldots, X_n] \leq X_n$$

- A stochastic process $X_1, X_2, X_3 \ldots$ is a submartingale if for all $n \geq 1$
$$\mathbf{E}[X_{n_1} \mid X_1, \ldots, X_n] \geq X_n$$

Drift Criteria

Consider a stochastic process $X_1, X_2, X_3 \dots$ such that for some $\alpha > 0$ and $J > 0$ we have

$$\mathbf{E}[X_{n+1} \mid X_1, \dots, X_n \text{ and } X_n > J] \leq -\alpha.$$

Under what conditions can we guarantee $\sup_n \mathbf{E}[X_n] < \infty$?

Example

$$\begin{aligned} \Pr(X_{n+1} = 1 \mid X_n = 0) &= 1 \\ \Pr(X_{n+1} = s+1 \mid X_n = s) &= 1 - \tfrac{1+\epsilon}{1+s} \\ \Pr(X_{n+1} = 0 \mid X_n = s) &= \tfrac{1+\epsilon}{1+s} \end{aligned}$$

$$\mathbf{E}[X_{n+1} \mid X_1, X_2, \dots, X_n \text{ and } J > 0] = (s+1)(1 - \frac{1+\epsilon}{1+s}) \leq -\epsilon$$

The process has a stationary distribution:

$$\pi(s) = (1 - \frac{1+\epsilon}{s})\pi(s-1) = \pi(0)\prod_{j \leq s}(1 - \frac{1+\epsilon}{j}) = \frac{C}{s^{1+\epsilon}}$$

In the stationary distribution $\mathbf{E}[X] = \sum_{s \geq 1} \frac{sC}{s^{1+\epsilon}} = \infty$.

Criteria for Ergodicity [Meyen & Tweedie]

Theorem (Foster's Criteria)

An aperiodic and irreducible Markov chain on a countable state space Ω *is positive recurrent (ergodic) if there is a function* $f : \Omega \to R^{+}$, *a finite set* $C \subset \Omega$, *and a constant* $\beta > 1$ *such that*

$$\begin{array}{lll} E[f(X_{t+1}) - f(X_t) \mid X_t = x] & \leq -\beta & x \notin C \\ E[f(X_{t+1}) \mid X_t = x] & \leq \infty & x \in C \end{array} \tag{1}$$

“Proof”

$$V = \max_{x \in C} f(x)$$

$$C' = \{x \mid f(x) \leq V\}$$

Assume that the chain is at $y \notin C'$. The expected time to return to C' is

$$O(f(y)) < \infty.$$

The set C', and thus the whole chain, are positive recurrent.

Application to Load Balancing

Let $f(X)$ be the total load in the system at state X.
Let the set C be all the states with total load bounded by $\Theta = O(n \log n)$.

$$C = \{X \mid f(X) \leq \Theta\}.$$

For $X \notin C$

$$E[f(X_{t+1}) - f(X_t) \mid X_t = x] = \lambda n \tau - (1 - \frac{1}{n^c})\Delta.$$

For $X \in C$

$$E[f(X_{t+1}) \mid X_t = x] \leq \Theta + \lambda n \tau < \infty$$

But Stationary Is Not What We Need...

- Consider a Markov chain on N with $\pi_i = \frac{c}{i^{1.5}}$.
- Let $f(i) = i$.
- The chain is ergodic by $E[f(i)]$ is not bounded.

Theorem (Foster's Criteria)

Given an aperiodic and irreducible Markov chain on a countable state space, and a function $f : \Omega \to [1, \infty]$*, the following two conditions are equivalent:*

1. *The chain is ergodic with stationary distribution* π*, and* $E_\pi[f(x)]$ *is finite;*
2. *There is a finite set* C *such that*

$$\sup_{x \in C} E[\sum_{i=1}^{T_{return}} f(X_i) \mid X_0 = x] < \infty,$$

where $T_{return} \geq 1$ *is time of first return to* C*.*

Application to the Load Balancing Algorithm

Let $f(X)$ be the total load in the system at state X.
Let the set C be all the states with total load bounded by $\Theta = O(n \log n)$.

$$C = \{X \mid f(X) \leq \Theta\}.$$

Length of Bad Segment

Total new load in a bad epoch is $\tau n = O(n \log n)$.
If the system is in a bad state, a successful epoch reduces the load by $\Delta = O(n \log n)$.
If the system is in a bad state during i epochs, at least αi epochs must be unsuccessful (for some constant $\alpha > 0$).

$$\begin{aligned} Pr(T_{return} \geq i) &\leq \binom{i}{\alpha(i-1)} \left(\frac{2}{n^c}\right)^{\alpha(i-1)} \\ &\leq n^{-\alpha(c-1)(i-1)} \end{aligned}$$

$$E[T_{return}^2 \mid X_0 = x \in C] = O(1).$$

Applying

$$E[T_{return}^2 \mid X_0 = x \in C] = O(1),$$

and

$$f(X_i) \leq T_{return}\tau n + \Theta,$$

$$\begin{aligned} & E[\sum_{i=1}^{T_{return}} f(X_i) \mid X_0 = x] \\ \leq \quad & E[\sum_{i=1}^{T_{return}} (\Theta + i\tau n)] \\ \leq \quad & E[(T_{return}^2 \tau n + T_{return}\Theta] < \infty. \end{aligned}$$

Do We Need the Markovian Property?

Stochastic Adversarial Queuing Model:

- Stochastic analysis subject to minimum restrictions on the set of possible input distributions.

Let

- $N(t, t + w) = \#$ of new jobs in the interval $[t, t + w]$.
- H_t - history till time t.

Definition

An adversary is a (w, ρ)- stochastic adversary if for all t,

$$E[N(t, t+w) \mid H_t] \leq \rho w.$$

Definition

A (w, ρ)-stochastic adversary is properly bounded if for some constants $p > 2$ and V, and for all t

$$E[(N(t, t+w))^p \mid H_t] \leq V.$$

Theorem (Pemantle & Rosenthal - 99)

Let $X_1, X_2, \ldots$ *be a sequence of nonnegative random variables satisfying the following conditions:*

1. *There exist positive constants* α *and* Θ *such that for all* $x_1, \ldots, x_i$ *with* $x_i > \Theta$, $\mathbf{E}[X_{i+1} - X_i | X_1 = x_1, \ldots, X_i = x_i] \leq -\alpha$.
2. *There exists a positive constant* Ψ *and a* $p > 2$ *such that for all* $x_1, \ldots, x_i$, $\mathbf{E}[|X_{i+1} - X_i|^p | X_1 = x_1, \ldots, X_i = x_i] \leq \Psi$.

Then there exist $\Psi = \Psi(X_0, \alpha, \Theta, \xi)$ *and* t_0 *such that for all* $t \geq t_0$,

$$\mathbf{E}[X_t | X_0] \leq \Psi + \max(0, X_0 - \Theta).$$

Furthermore, assuming that p is a constant, then

$$\Psi = O\left(\Theta + \alpha\left(1 + \frac{\xi}{\alpha^{p}}\right)^{3p}\right),$$

as $n \to \infty$.

Back to Load Balancing

More general input distribution, no Markovian condition:
$E[\ \#\text{ of new jobs in } w \text{ steps}] < n \cdot w$
$E[\#\text{ of new jobs in } w \text{ steps}^p] < M$, for some $p > 2$

Theorem

For any (w, ρ)- properly bounded stochastic adversary with $\rho < 1$,

- *The load-balancing protocol is stable;*
- *Limit expected load in the system is $O(wn \log n)$.*

Proof

Define an epoch as $\max[\tau, w]$ time steps.
$Y_i =$ be the total load in the system at the beginning of segment i.
When load $> \Theta$, a successful epoch reduces the load by at least Δ.

Lemma

- *There exist positive constants* Δ', *such that for all* $y_1, \ldots, y_i$ *with* $y_i > \Theta$,

$$E[Y_{i+1} - Y_i | Y_1 = y_1, \ldots, Y_i = y_i] \leq -\Delta'$$

- *There exists a positive constant* $\Psi = O(wn \log n)$ *and a* $p > 2$ *such that for all* $y_1, \ldots, y_i$,

$$E[|Y_{i+1} - Y_i|^p | Y_1 = y_1, \ldots, Y_i = y_i] \leq \Psi.$$