Modeling and analysis of placement group heuristics

Asser Tantawi

IBM TJ Watson Research Center

Concepts

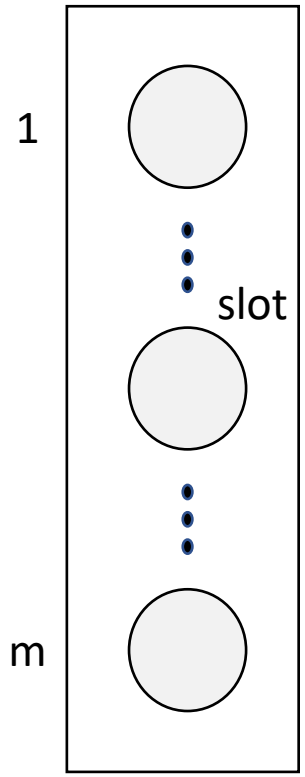
- System
  - multiple racks
- Rack
  - capacity for m (homogeneous) VMs (slots)
- Placement Group
  - size n (homogeneous) VMs
  - placed with rack affinity
  - VM requests arrive over a period of time
- VM
  - single and group arrivals
  - have a lifetime, thus utilization (allocation) load on the system

Question

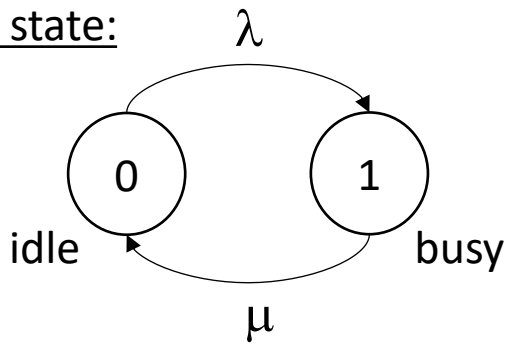
- Given a **heuristic for selecting a rack** to place a group
- what is the impact of
  - heuristic
  - system load
  - group size (relative to rack capacity)
  - group requests arrival period
- on placement success
  - probability of being able to place all members of the group

Modeling

rack



slot state:



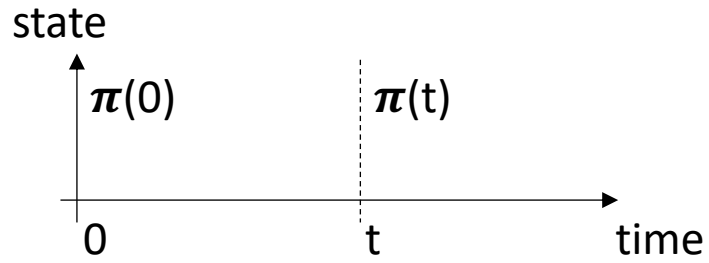
rate transition matrix:

$$\mathbf{Q} = \begin{bmatrix} -\lambda & \lambda \\ \mu & -\mu \end{bmatrix}$$

utilization:

$$\rho \triangleq \text{Prob}[busy] = \frac{\lambda}{\lambda + \mu}$$

transient analysis:



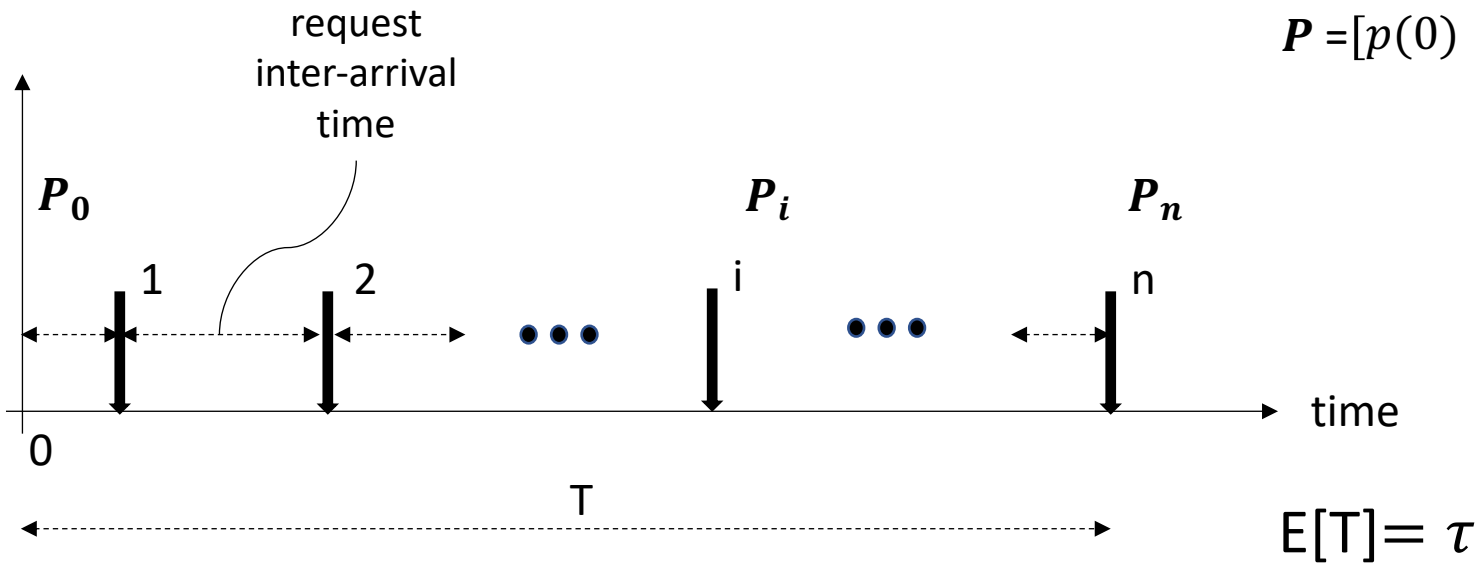
$$\boldsymbol{\pi}(t) = \boldsymbol{\pi}(0)\, e^{\boldsymbol{Q}t}$$

$$\boldsymbol{\pi} = \begin{bmatrix} \pi_0 & \pi_1 \end{bmatrix}$$

Modeling

$\boldsymbol{P} = [p(0) \quad p(1)]$

group arrival process:



$\mathrm{E}[\mathrm{T}] = \tau$

Assuming Poisson request arrivals with rate θ

$$\boldsymbol{P}_i = \boldsymbol{P}_0 \, \theta^i [(\boldsymbol{I}\theta - \boldsymbol{Q})^{-1}]^i$$

Assuming independence, group success probability is

$$\nu = \prod_{i=1}^{n} 1 - p_i(1)^{m-i+1}$$

success probability

Color code	Gap
0.98 - 1.00	0.02
0.94 - 0.98	0.04
0.86 - 0.94	0.08
0.70 - 0.86	0.16
0.38 - 0.70	0.32
0.00 - 0.38	

heavy rack (bad heuristic)

medium rack

light rack (good heuristic)

rho = 0.4

low load

		p00 = 0.2					p00 = 0.4					p00 = 0.8				
	1/2	0.673	0.757	0.860	0.944	0.984	0.981	0.986	0.991	0.995	0.998	1.000	1.000	1.000	1.000	1.000
	1/4	0.902	0.935	0.967	0.988	0.997	0.998	0.999	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
n/m	1/8	0.965	0.979	0.991	0.997	0.999	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	1/16	0.987	0.993	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		1/16	1/8	1/4	1/2	1	1/16	1/8	1/4	1/2	1	1/16	1/8	1/4	1/2	1
				tau					tau					tau		

rho = 0.6

medium load

		p00 = 0.2					p00 = 0.4					p00 = 0.8				
	1/2	0.644	0.710	0.797	0.881	0.937	0.975	0.975	0.975	0.975	0.975	1.000	1.000	1.000	0.998	0.992
	1/4	0.890	0.917	0.947	0.972	0.987	0.997	0.997	0.997	0.997	0.997	1.000	1.000	1.000	1.000	1.000
n/m	1/8	0.959	0.971	0.983	0.992	0.996	0.999	0.999	0.999	0.999	0.999	1.000	1.000	1.000	1.000	1.000
	1/16	0.984	0.989	0.994	0.997	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		1/16	1/8	1/4	1/2	1	1/16	1/8	1/4	1/2	1	1/16	1/8	1/4	1/2	1
				tau					tau					tau		

rho = 0.8

high load

large group →

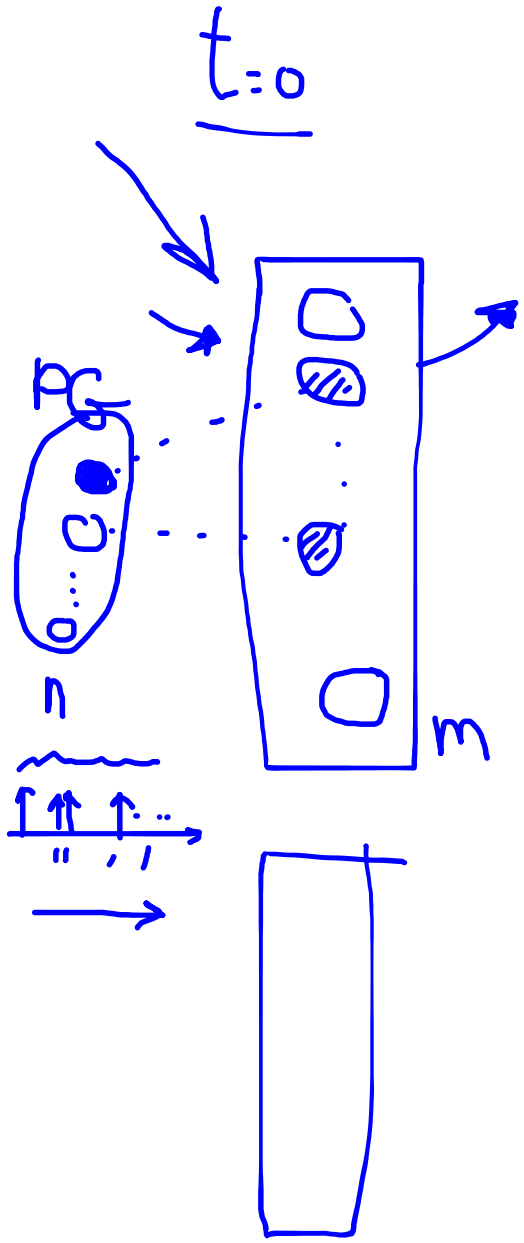
		p00 = 0.2					p00 = 0.4					p00 = 0.8				
	1/2	0.557	0.557	0.557	0.557	0.557	0.950	0.918	0.849	0.742	0.647	1.000	0.998	0.982	0.889	0.731
	1/4	0.847	0.847	0.847	0.847	0.847	0.993	0.986	0.970	0.938	0.902	1.000	1.000	0.999	0.984	0.943
n/m	1/8	0.938	0.938	0.938	0.938	0.938	0.998	0.996	0.990	0.979	0.966	1.000	1.000	1.000	0.997	0.986
	1/16	0.972	0.972	0.972	0.972	0.972	0.999	0.998	0.996	0.991	0.986	1.000	1.000	1.000	0.999	0.997
		1/16	1/8	1/4	1/2	1	1/16	1/8	1/4	1/2	1	1/16	1/8	1/4	1/2	1
				tau					tau					tau		

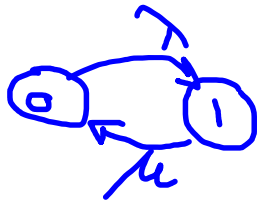
← small group

fast arrivals

slow arrivals

backup

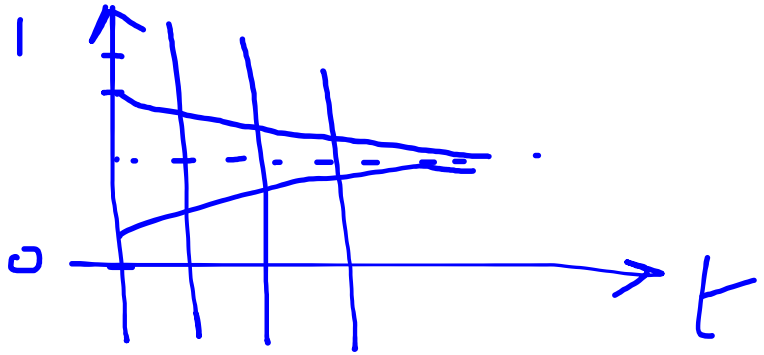




$\rho = \frac{\lambda}{\mu} < 1$

$P_0 = 1 - \rho$

$P_i = \rho^i$



① S

② n ? m

③ PG arrival rate