Outline

1. Tools from Operations Research
   - Little's Law (average values)
   - Unreliable Machine(s) (operation dependent)
   - Buffers (zero buffers & infinite buffers)
   - M/M/1 Queue (effects of variation)
2. Applications
   - See other hand outs…

Little's Law

$$N = \lambda\, T$$

N = Average parts in the system

λ = Average arrival rate

T = Average time in the system

Ref. L. Kleinrock, “Queueing System, Vol 1 Theory, Wiley, 1975

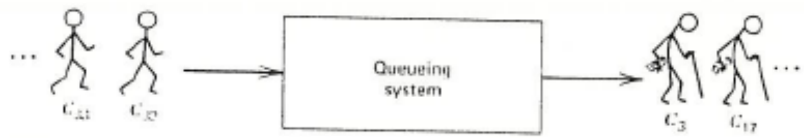
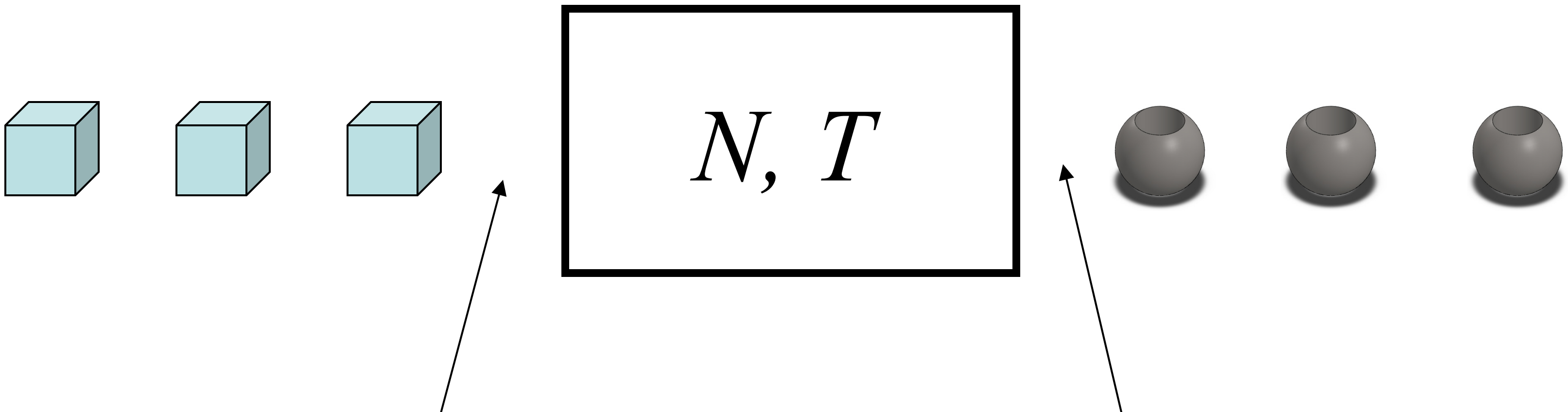


Figure 2.1 A general queueing system.

Queueing Systems



Observer here
Observing arrivals;
$\alpha(t)$ = number of arrivals in (0,t)

Observing departures;
$\delta(t)$ = number of departures in (0,t)

$$N(t) = a(t) - d(t)$$

Number in the system, parts or customers

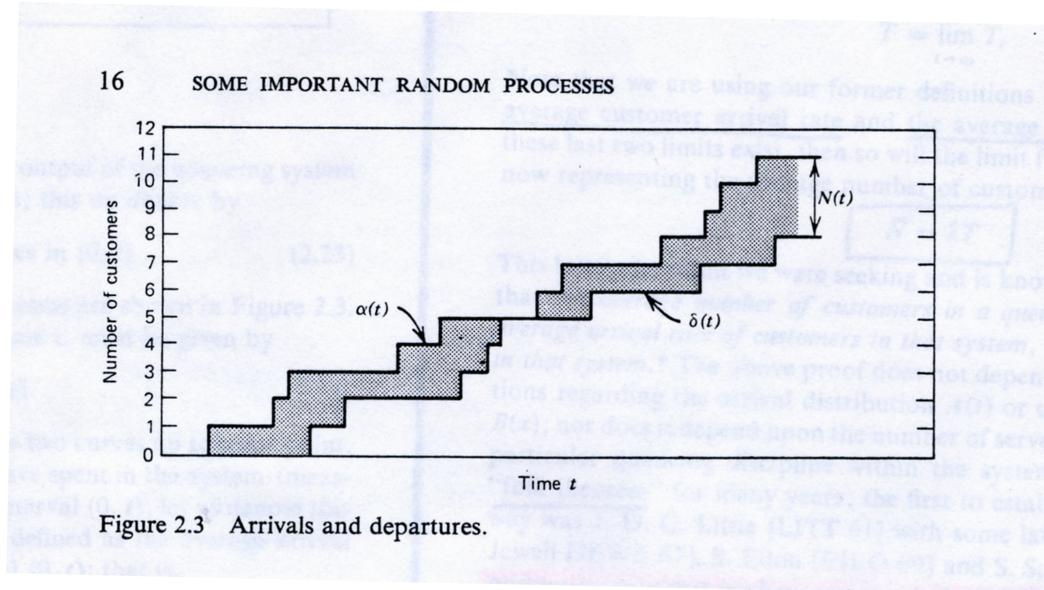


Figure 2.3 Arrivals and departures.

Ref. Kleinrock Vol 1, 1975 See p.16, 17

$$N(t) = a(t) - d(t)$$

$$g(t) = \int_0^t N(t)\,dt$$

$g(t) =$ customer -seconds (shaded area in figure)

average number of customers in the system $= \bar{N} = \dfrac{g(t)}{t}$

average arrival rate $= l_t = \dfrac{a(t)}{t}$

average time per customer $= T_t = \dfrac{g(t)}{a(t)} = \dfrac{\bar{N}}{l_t}$

$$T = \lim_{t \to \infty} T_t$$

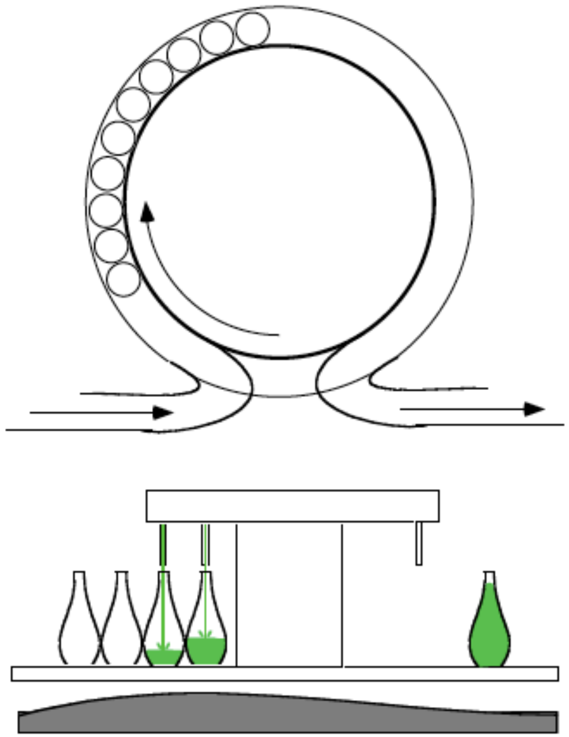
$$l = \lim_{t \to \infty} l_t$$

this gives $\bar{N} = l_t \cdot T_t$

assuming the limits exist gives $\bar{N} = l \cdot T$ (or $L = \lambda W$)

$$\bar{N} = l \cdot T$$

Typical Dial Machine



10/23/13

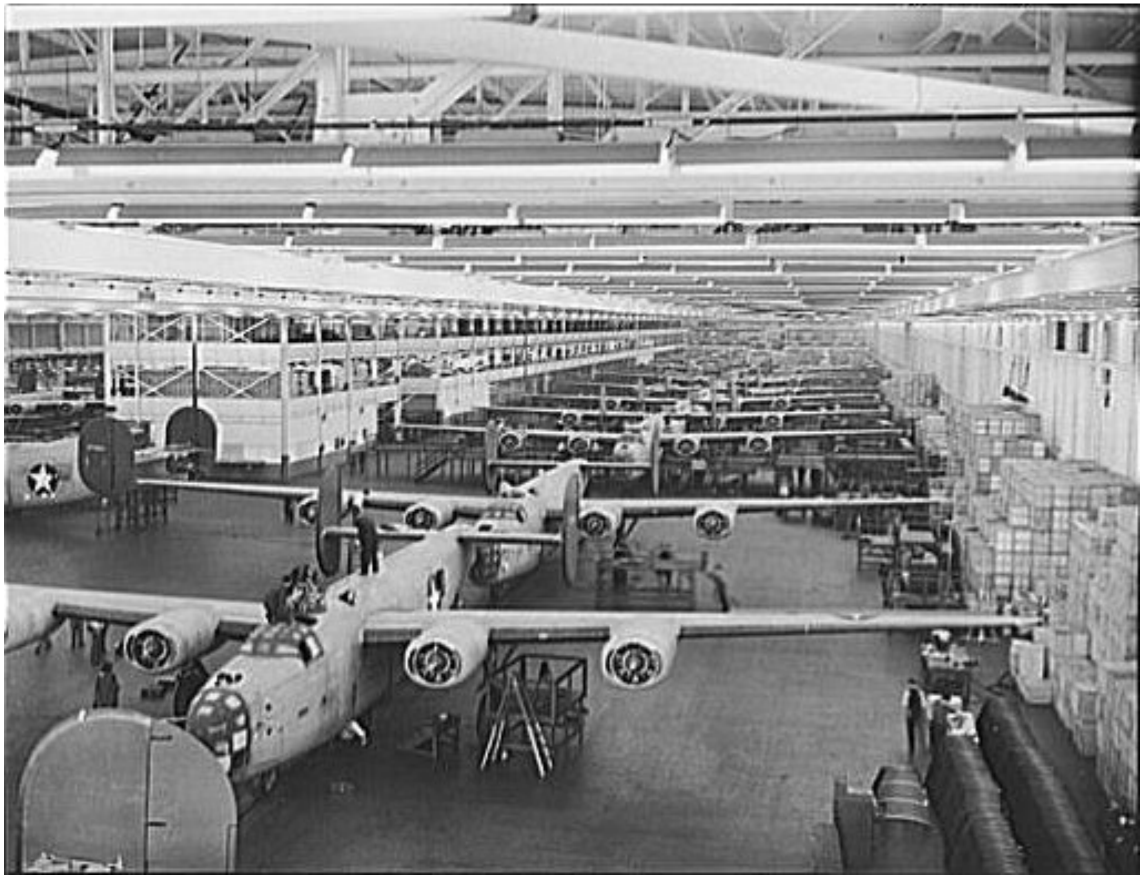
Q. You want a high rate of production λ*, but if you fill too fast the liquid comes out. What do you do?*

A. Fill while the bottle is moving making T *long enough to avoid losing any liquid.*

This results in long lines and large factories $\bar{N} = l \cdot T$

Ford’s Willow Run Factory

Moving assembly line production of B-24s



Ford’s Willow Run plant - 10 mo delay, but in 1944 produced 453 airplanes in 468 hrs

About 1 plane every hour!

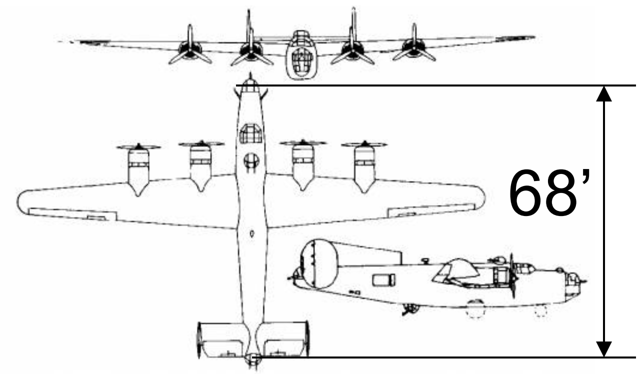
How long did they work on assembly?

- Production rate when fully running was about 1 plane every hour
- Little's Law: $L = \lambda W$
- $\lambda = 1$ plane/hr
- L = ? "Assembly line was over one mile"
- W = ?

How long did they work on assembly?

- Production rate when fully running was about 1 plane very hour
- Little's Law: $L = \lambda W$
- λ = 1 plane/hr
- L = 5280'/68' = 78 planes,

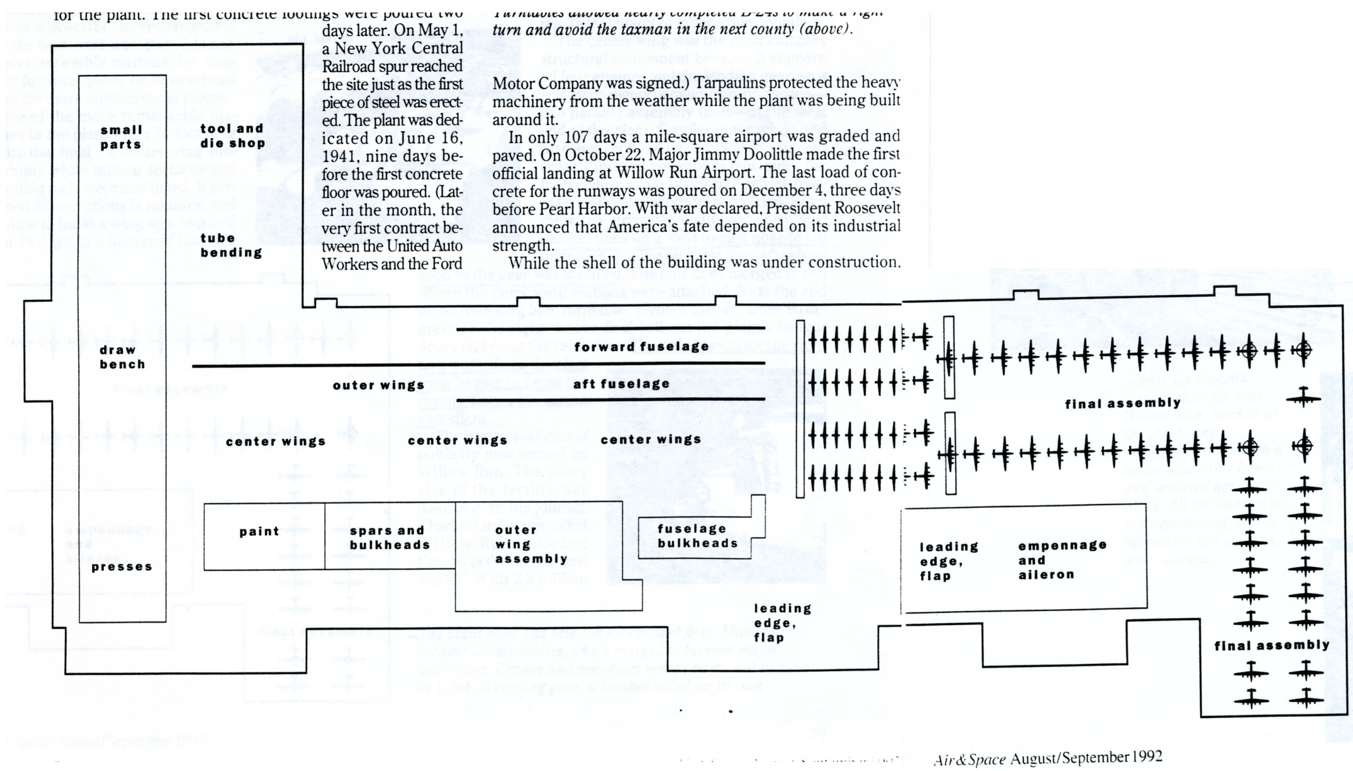
  (if heel to toe for one mile)
- $W = L/\lambda \approx$ 78 hours



Willow Run

Two lines converge into one

Ford's Willow Run Factory

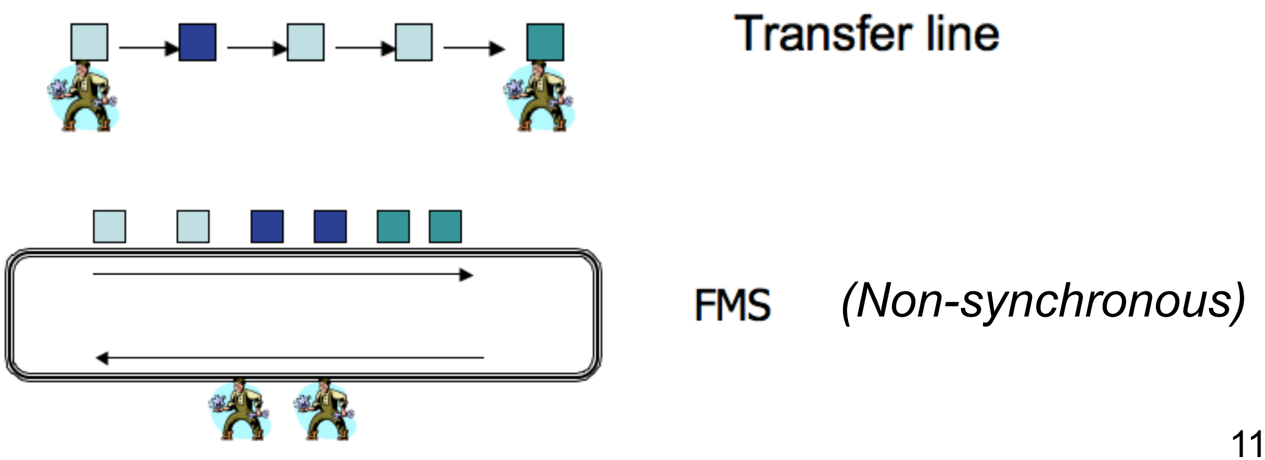


Assembly Line, L ~ 81 planes, implies around 81 hrs/plane

Ref; Air & Space Aug/Sept 1992

Applying Little's Law

- Boundaries are arbitrary, but you must specify eg. waiting time + service time
- Internal details are not considered eg. first in first out, flow patterns etc..



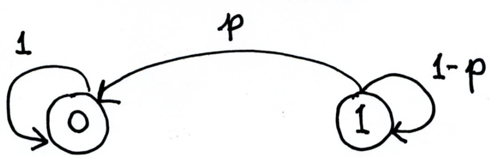
Unreliable Machine

- Ref S. B. Gershwin (Ch 2 of his book)
- Preliminaries: conditional probability and Markov chains - transition probabilities
- Discrete or continuous time - ODQ
- Probability machine is down - exponential distribution

Failure distribution

Discrete time model

State	Meaning
0	Machine down
1	Machine up



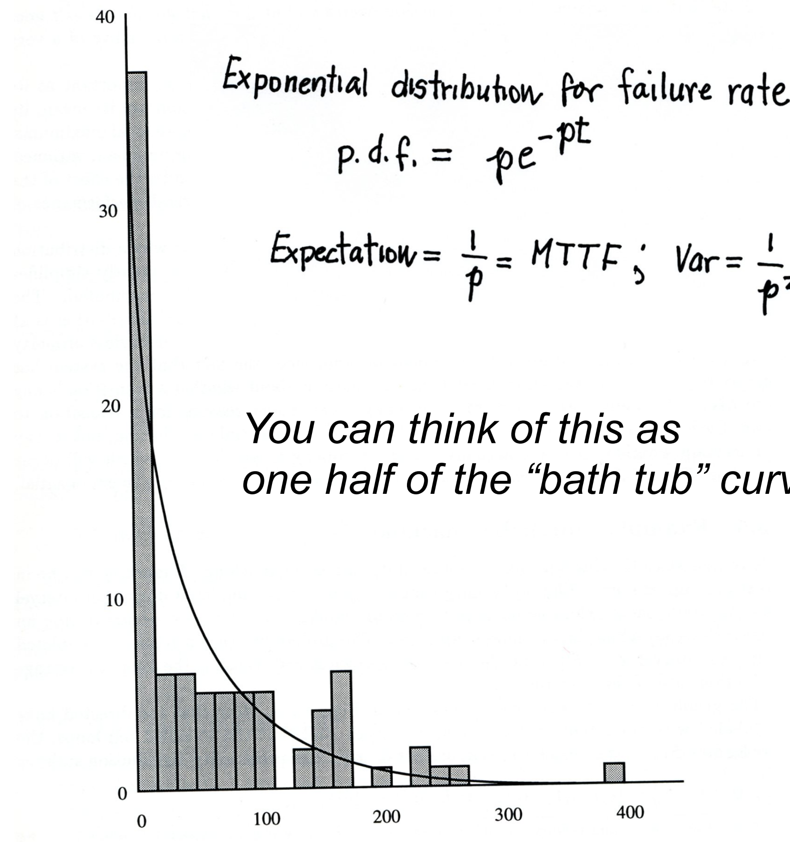
$$\begin{Bmatrix} P(0,t+1) \\ P(1,t+1) \end{Bmatrix} = \begin{bmatrix} 1 & p \\ 0 & 1-p \end{bmatrix} \begin{Bmatrix} P(0,t) \\ P(1,t) \end{Bmatrix}$$

transition matrix

Probability machine fails at time t = $p(1-p)^{t-1}$

Geometric distribution

Continuous time model

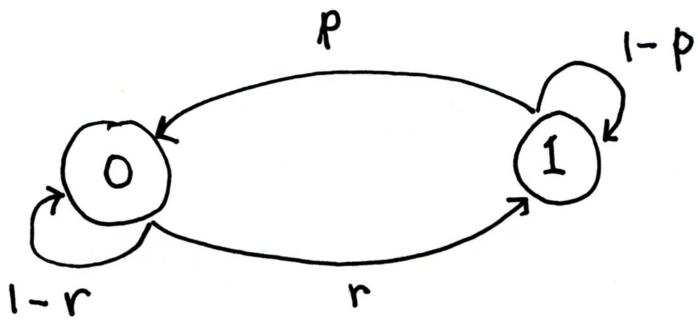


You can think of this as one half of the "bath tub" curve

Figure 2.7: Exponential Density Function and Samples

Note: MTTF = mean time to failure

Unreliable Machine with Repair



$$\begin{Bmatrix} p(0,t+1) \\ p(1,t+1) \end{Bmatrix} = \begin{bmatrix} 1-r & p \\ r & 1-p \end{bmatrix} \begin{Bmatrix} p(0,t) \\ p(1,t) \end{Bmatrix}$$

$$\frac{1}{p} = \text{MTTF} \quad ; \quad \frac{1}{r} = \text{MTTR}$$

Note: MTTR = mean time to repair

Discrete time:

Steady state solution for probability that machine is up =

$$\frac{\text{MTTF}}{\text{MTTF+MTTR}}$$

Continuous time

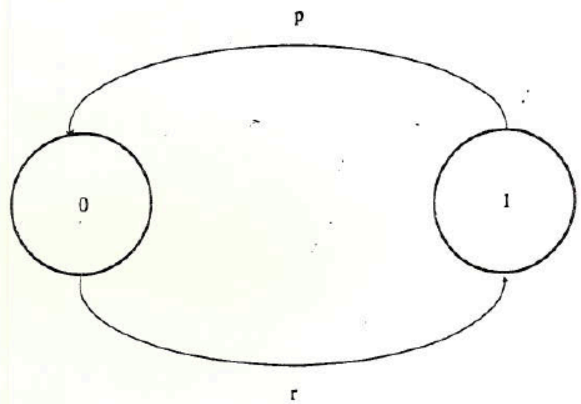


Figure 2.8: Graph of Markov Chain for Continuous Time Unreliable Machine Model

$$\frac{d\mathbf{p}(0,t)}{dt} = -\mathbf{p}(0,t)r + \mathbf{p}(1,t)p$$

$$\frac{d\mathbf{p}(1,t)}{dt} = \mathbf{p}(0,t)r - \mathbf{p}(1,t)p.$$

$$\boldsymbol{p}(0,t)+\boldsymbol{p}(1,t)=1$$

Continuous time

The solution is

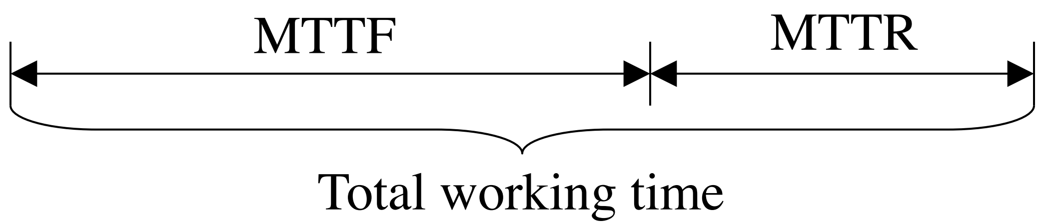
$$
\begin{aligned}
\mathbf{p}(0,t) &= \frac{p}{r+p} + \left[\mathbf{p}(0,0) - \frac{p}{r+p}\right] e^{-(r+p)t} \\
\mathbf{p}(1,t) &= 1 - \mathbf{p}(0,t).
\end{aligned}
$$

As $t \to \infty$, we have

$$\mathbf{p}(0) = \frac{p}{r+p}; \mathbf{p}(1) = \frac{r}{r+p}.$$

The average production rate is $\mathbf{p}(1)\mu$ or $\frac{r\mu}{r+p}$.

Single unreliable machine



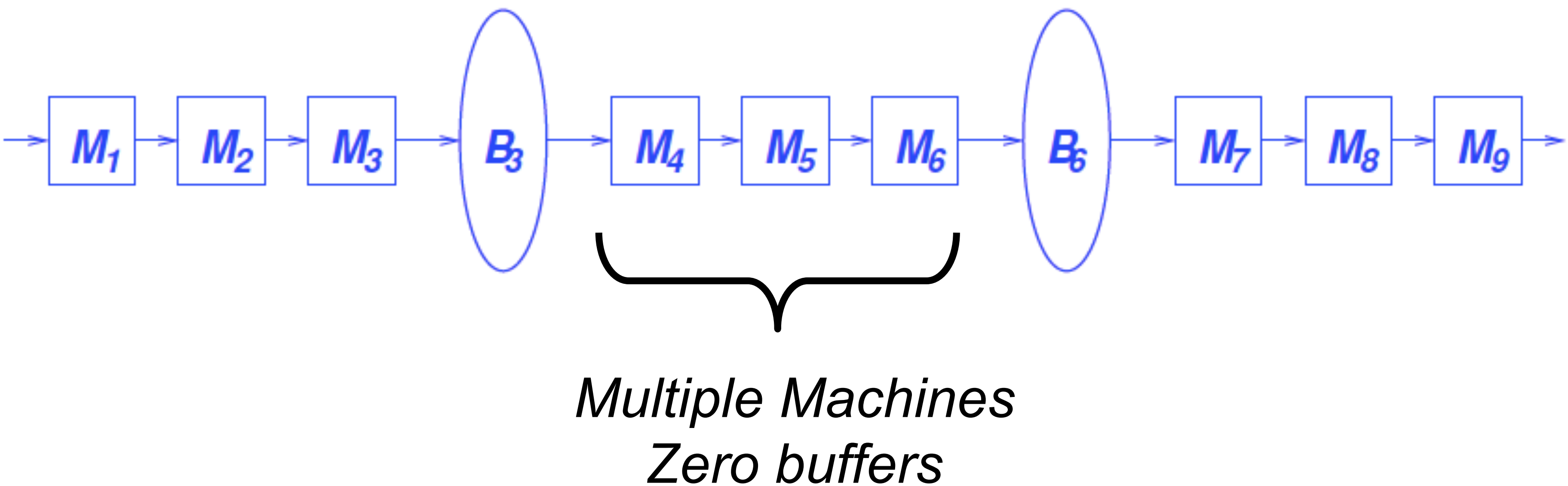
$$\text{Machine up} = \frac{\text{MTTF}}{\text{MTTF} + \text{MTTR}}$$

$$\text{Machine down} = \frac{\text{MTTR}}{\text{MTTF} + \text{MTTR}}$$

$$\text{Average Production rate} = \frac{1}{\tau} \times \frac{\text{MTTF}}{\text{MTTF} + \text{MTTR}}$$

Where, τ = operation time = $1/\mu$

Operational Dependent Failures



<u>Operation</u> dependent

e.i. machine can only fail when it is operating

Multiple Machine Case: Zero Buffers

Consider a long time interval T,

say there are m_i failures for machine i

$$\therefore \text{ Total downtime } = D = \sum_{i=1}^{k} m_i \, MTTR_i = \sum_{i=1}^{k} \frac{m_i}{r_i}$$

$$\text{Total up time } \quad U = T - D$$

$$\#\text{ failures} = m_i = \frac{U}{MTTF_i} = p_i U$$

this gives

$$U = T - U \sum_{i=1}^{k} \frac{p_i}{r_i}$$

or

$$\frac{U}{T} = \frac{1}{1 + \sum_{1}^{k} \frac{p_i}{r_i}}$$

for one machine

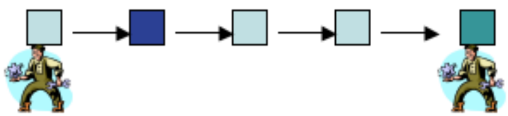
$$\frac{U}{T} = \frac{r_i}{r_i + p_i} = \frac{MTTF_i}{MTTF_i + MTTR_i}$$

Operation dependent

e.i. machine can only fail when it is operating

Unreliable Machine(s) Result

- Multiple identical machines (Transfer line)



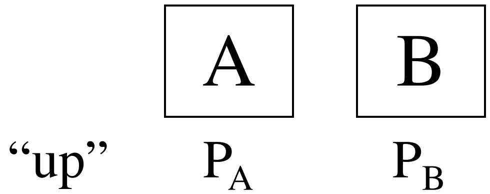
Buzacott's formula, $\mu = \frac{1}{\tau} \times \frac{1}{1+\sum_{1}^{k}\frac{\text{MTTR}}{\text{MTTF}}}$

- Single Machine

$$\mu = \frac{1}{\tau} \times \frac{\text{MTTF}}{\text{MTTF} + \text{MTTR}}$$

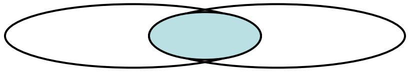
τ = *service time without failures*

Time Dependent Estimation of μ



Assumption: time dependent failure
(A and B are two processes with nominal rate $\mu=1/\tau$ in series. Their behaviors are not dependent on each other.)

Probability that both A and B are up is $A \cap B$



$$A \cap B = P_A P_B$$

$$\text{Production rate} = \boxed{\frac{1}{\tau} P_A P_B} = \frac{1}{\tau} \frac{MTTF_A}{MTTF_A + MTTR_A} \times \frac{MTTF_B}{MTTF_B + MTTR_B}$$

$$= \frac{1}{\tau} \frac{1}{1+\alpha_A} \times \frac{1}{1+\alpha_B} \qquad \text{Where, } \alpha_i = \frac{MTTR_i}{MTTF_i}$$

$$= \frac{1}{\tau} \frac{1}{1+\alpha_A+\alpha_B+\alpha_A \alpha_B}$$

Estimation of μ (continued)

$$= \frac{1}{\tau} \frac{1}{1 + \alpha_A + \alpha_B + \alpha_A \alpha_B}$$

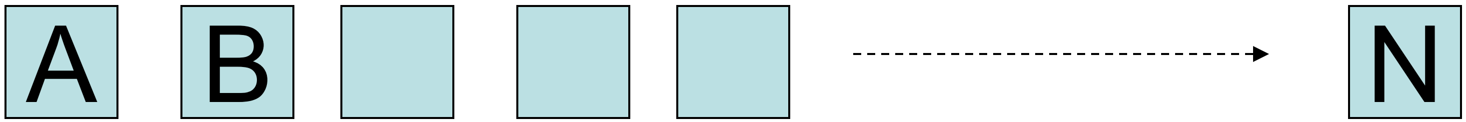
Note: $\alpha_A \alpha_B << 1$

Ignoring higher order terms,
Same as Buzacott's result

$$\mu \approx \frac{1}{\tau} \frac{1}{1 + \sum_{1}^{2} \alpha_i}$$

Note: seems to give the same answer as Buzacott,
but second order terms can become important for large systems.
Need to differentiate between <u>operation</u> and <u>time dependent</u> failures

Example:Transfer Line



infinite buffer $\mu_0 = (1/\tau \times p)_{bottleneck}$

zero buffer $\mu_\infty = 1/\tau \times p_A p_B \ldots p_N$

example; transfer line, all p = 0.9

$\mu = (0.9)^N \times 1/\tau$

N=1 $\mu = .9 \times 1/\tau$

N=10 $\mu = .35 \times 1/\tau$

N=100 $\mu = .00003 \times 1/\tau$

Time dependent

$\mu = (1/(1+0.111N)) \times 1/\tau$

N=1 $\mu = .9 \times 1/\tau$

N=10 $\mu = .47 \times 1/\tau$

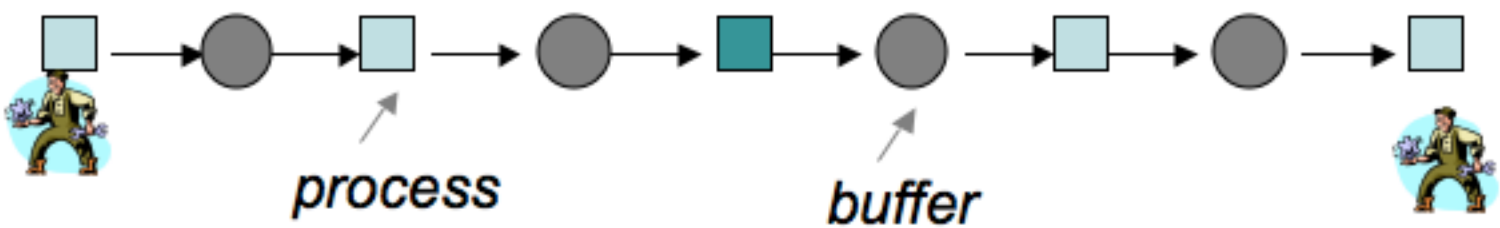
N=100 $\mu = .0825 \times 1/\tau$

Operation dependent

Summary: Production Rates

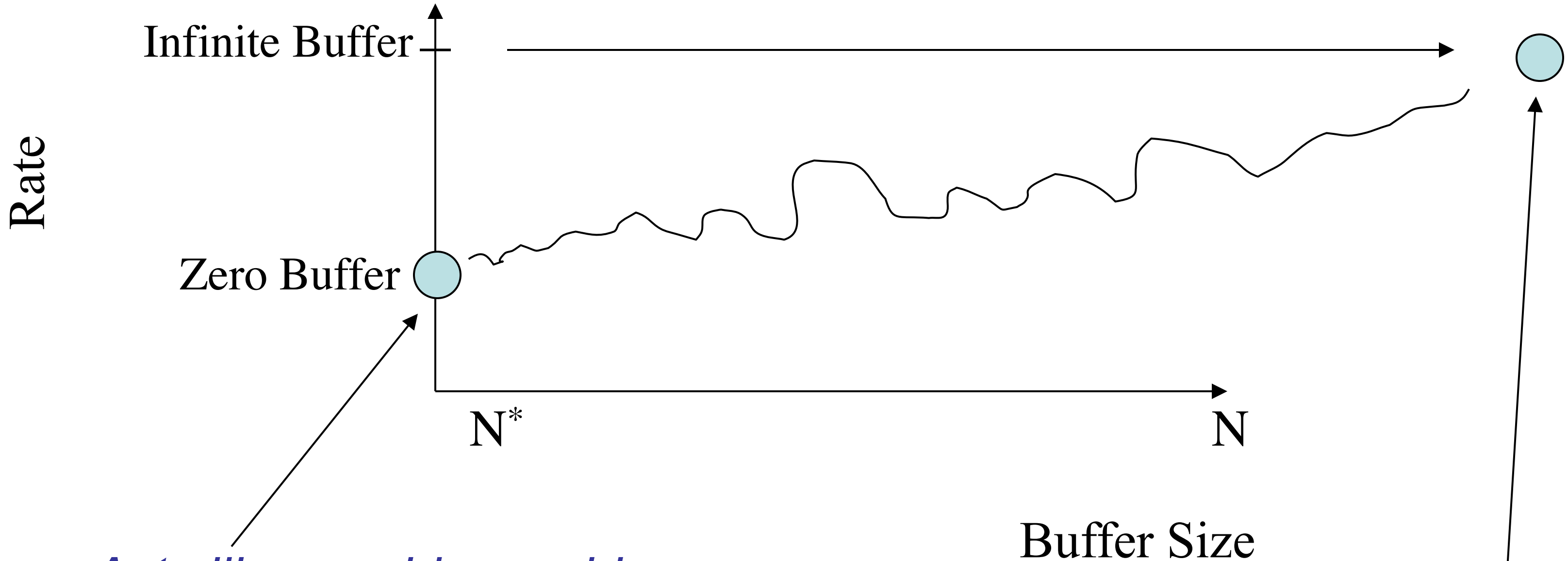
Zero Buffer: $$\frac{1}{t} \cdot \frac{1}{1 + \sum_{1}^{n} \frac{MTTR_i}{MTTF_i}}$$ *Transfer line*

Infinite Buffer: $$\min\left(\frac{1}{t_i} \cdot \frac{MTTF_i}{MTTF_i + MTTR_i}\right)$$ *Bottleneck*



Finite Buffer Size

How do the two cases connect for finite buffers?



Acts like one big machine

Machines are independent rate is controlled by the slowest machine "bottleneck"

A small amount of buffer space helps a lot, but too much is costly

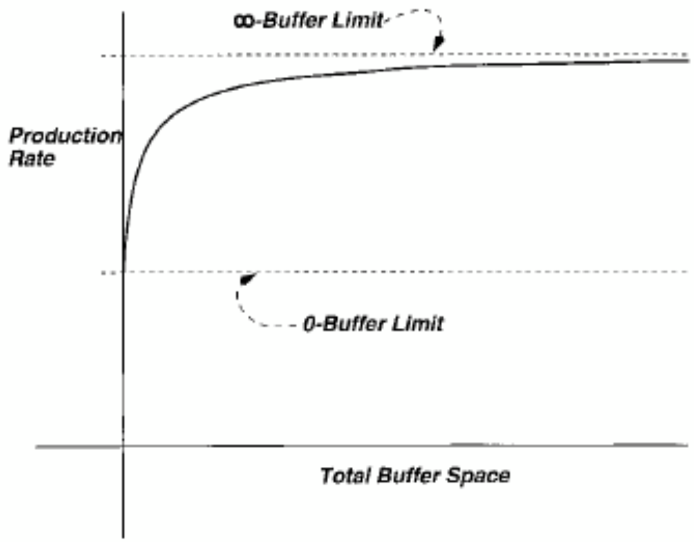
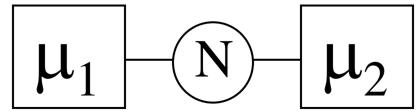


Figure 3: The production rate increases as in-process inventory space increases. This increase is rapid at first and then small. The upper and lower limits are easy to calculate, but the rest of the curve requires the decomposition method.

Finite buffer approximation

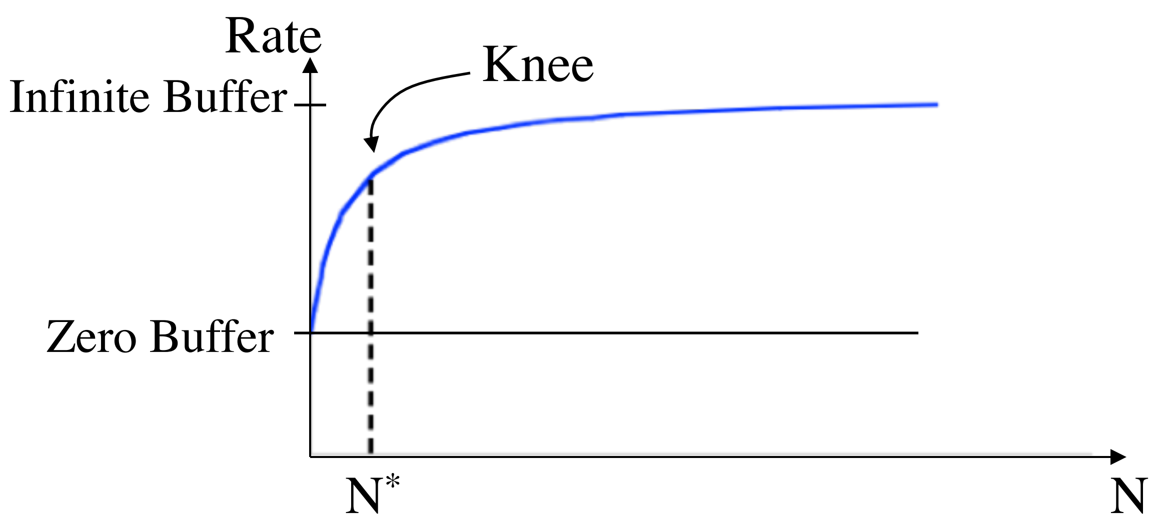
For a two machine system :



Average Downtime is $\dfrac{MTTR_1 + MTTR_2}{2}$

and, $\mu 1 \approx \mu 2$, call the rate μ.

Gershwin's Approximation: $N^* \approx 2 \text{ to } 6 \times \overline{MTTR} \times \mu$



Simulation of a 20 machine, 19 buffer (cap = 10 parts) Transfer line. Each machine with one minute cycle time could produce 4800 parts per week. MTTF 3880 minutes, MTTR 120 minutes. *See Gershwin p63-64*

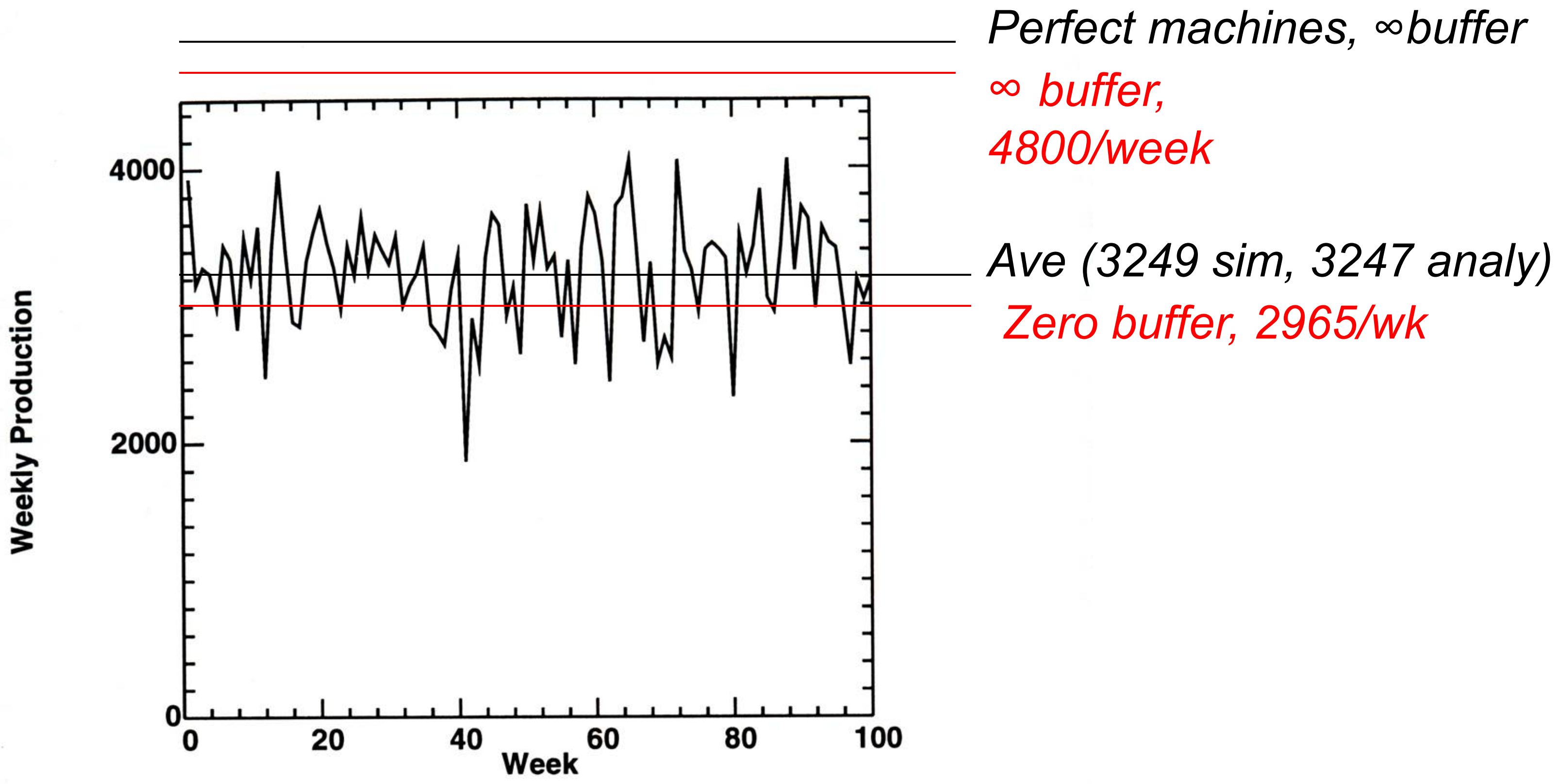
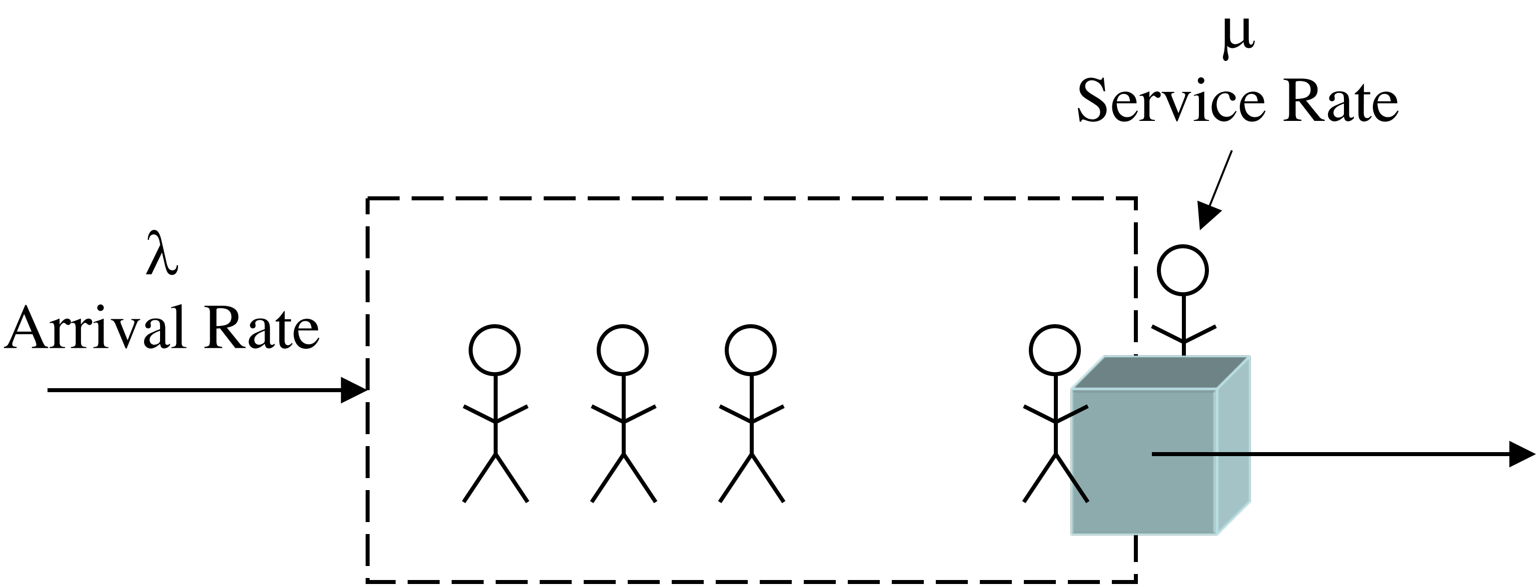


Figure 3.2: Production Variability

M/M/1 Queue



..how the inventory in the system grows as you approach capacity

(λ & μ vary according to exponential distribution)

Steady State $(\lambda < \mu)$

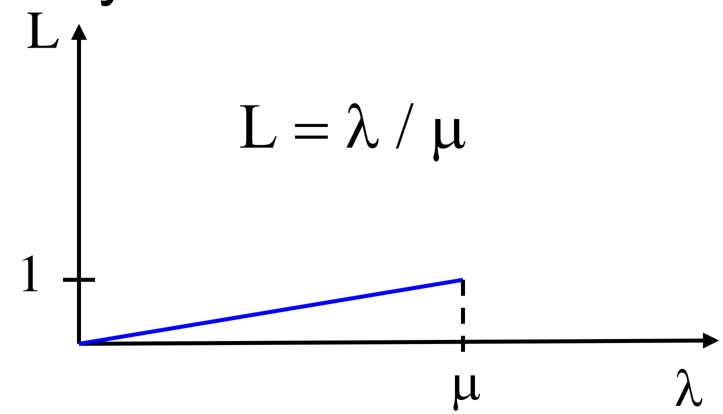
$\lambda < \mu$

Consider the <u>deterministic</u> case:

- How many people are in the system?

A.

$\lambda = 0$	L = 0
$0 < \lambda < \mu$	0 < L < 1
$\lambda = \mu$	L = 1



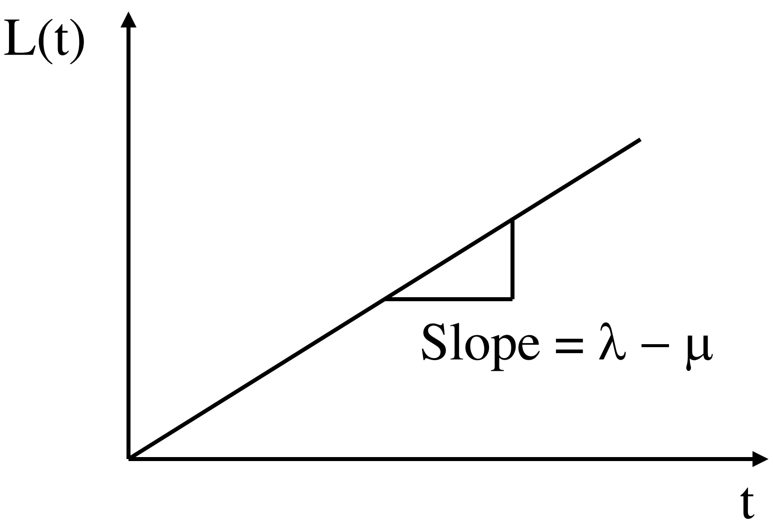
Note: From Little's Law : Time in system, $W = L / \lambda$

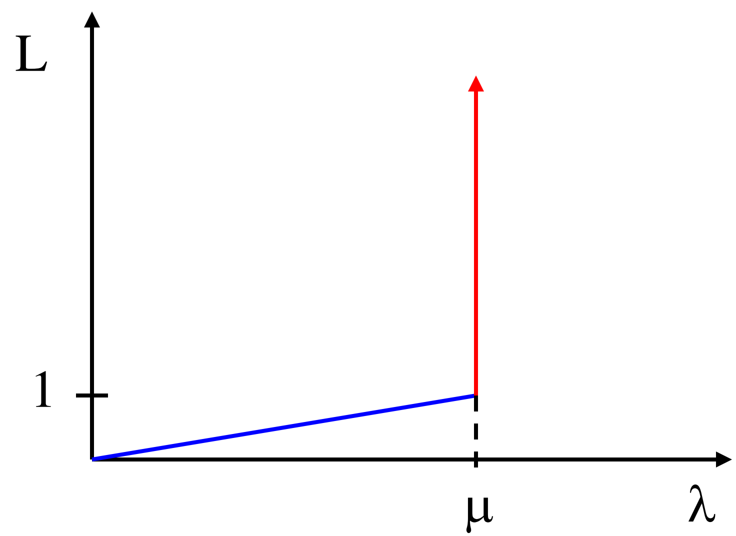
Since $L = \lambda / \mu$ for $\lambda < \mu$ ➔ $W = 1 / \mu$

When $\lambda > \mu$

- What happens at $\lambda > \mu$?
- There is no steady state, parts in the system grow without limit.

As $t \to \infty$, $L \to \infty$



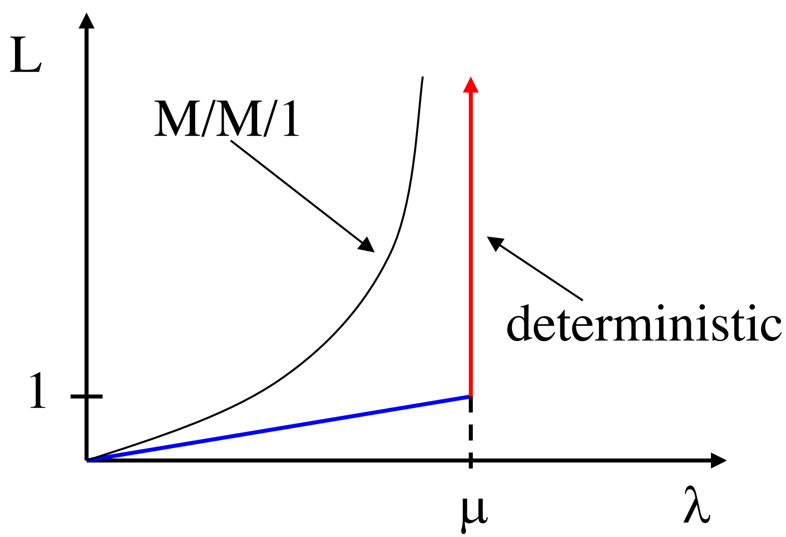


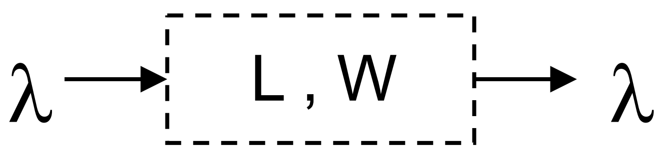
M/M/1 Queue Result

Arrival rate = λ , Service rate = μ, where $\lambda \leq \mu$

$L =$ "Inventory" $= \lambda / (\mu - \lambda)$

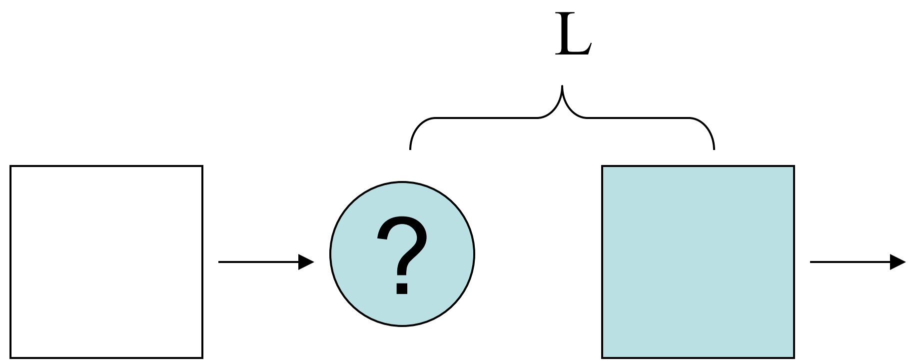
$W =$ Time in system $= 1 / (\mu - \lambda)$





See Notes: Principles needed to derive M/M/1 queue result - on website

example: two processes



Process A:
never starved
outputs parts
at average rate λ
with an exponential distribution

Process B:
with average process rate $\mu = (5/4)\lambda$ also
with an exponential distribution

Parts in the system: deterministic: $L = 4/5$; M/M/1: $L = 4$

M/M/1 Queue interpretation

- Overly simplistic but tractable
- Arrivals (always “on”) vs departures (stop when the queue is empty)
- Show behavior as you approach capacity

G/G/1 Queue result

A more useful queueing result is for the G/G/1 queue

G ⇒ general distributions for arrival and service times with

$$\text{Expectation (arrival)} = \frac{1}{\lambda} \; ; \; \text{Exp(service)} = \frac{1}{\mu}$$

$$(\text{Coef of variation})^2 = \frac{\text{Variance}}{\text{mean}^2} \Rightarrow c_\lambda^2 \text{ and } c_\mu^2$$

$$W_q = \text{Time in queue (approx)} = \left(\frac{c_\lambda^2 + c_\mu^2}{2}\right)\left(\frac{u}{1-u}\right)\left(\frac{1}{\mu}\right)$$

$$u = \text{utilization} = \frac{\lambda}{\mu}$$

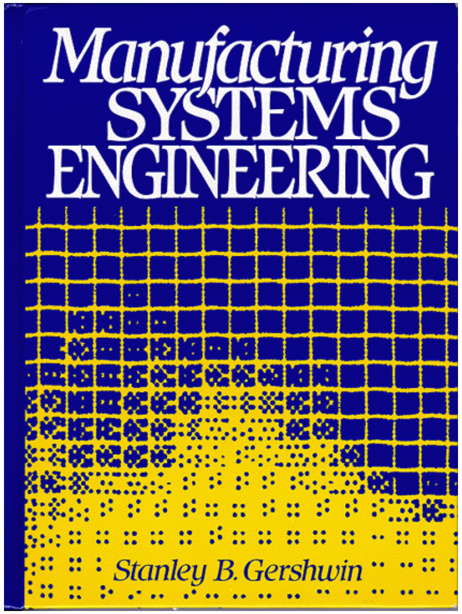
Limitations: $c_\lambda^2, c_\mu^2 \le 1$; $\frac{\lambda}{\mu} < 0.95$

Ref. Hopp & Spearman, Factory Physics p. 277
(this approximation used in several commercially available mfg queueing analysis packages)

Note: $W = W_q + 1/\mu$

Note: this result shows the same nonlinear rise in W and L as the system approaches capacity as the M/M/1 queue did

For more details take 2.854



Optional references:

1. *Kleinrock (Little's Law)- handout*
2. *Gershwin, Mfg Systems Engineering, Ch 2 & 3*
3. *Gershwin,_Notes on…(on our website, covers math for exponential distribution, unreliable machine and M/M/1 Queue)*