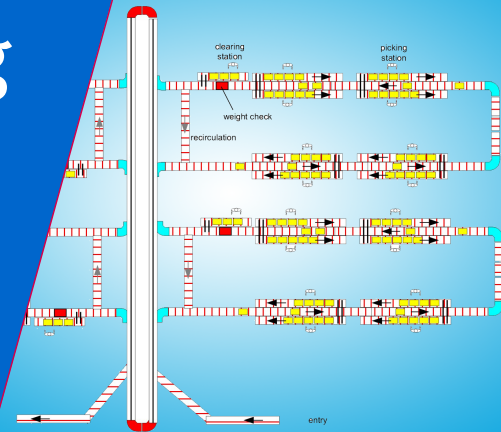


Analysis of Manufacturing Systems 4AB00

Ivo Adan



TU/e

Technische Universiteit
Eindhoven
University of Technology

Lecture notes on Analysis of Manufacturing Systems

- Chapter 3: 3.1, 3.2, 3.3

- **Effective process time** is total time seen by a job at a station
- Coefficient of variation c is **relative measure** of variability

$$c = \sigma/t$$

where t is the mean and σ the standard deviation

- **Preemptive breakdowns:**

$$A = \frac{m_f}{m_f + m_r} \quad (\text{availability})$$

$$t_e = \frac{t_0}{A}$$

$$c_e^2 = c_0^2 + A(1 - A)\frac{m_r}{t_0} + c_r^2 A(1 - A)\frac{m_r}{t_0}$$

where

- t_0 and c_0 are mean and cv of **natural** process time
- m_f is mean **exponential** time to failure
- m_r and σ_r are mean and standard deviation of time to repair

Recap: Process time variability

4/32

Example: What is better: short frequent stops or long rare ones?

Suppose $t_0 = \sigma_0 = 10$ minutes

Machine M_1 : $m_f = 90, m_r = 10, \sigma_r = 0$

Machine M_2 : $m_f = 900, m_r = 100, \sigma_r = 0$

Both machines have same availability $A = 0.9$, so **same effective capacity**

$$r_e = \frac{1}{t_e} = \frac{A}{t_0} = \frac{0.9}{10} = 0.09$$

but $c_e^2 = 1.09$ for M_1 and $c_e^2 = 1.9$ for M_2 !

- **Flow** refers to transfer of jobs from one station to another
- t_a and σ_a are mean and standard deviation of time between arrivals
- Arrival rate

$$r_a = \frac{1}{t_a}$$

- Coefficient of variation of time between arrivals

$$c_a = \frac{\sigma_a}{t_a}$$



Low c_a arrivals



High c_a arrivals

- Times between arrivals are independent and **exponential** with rate λ
- Memoryless property

$$P(\text{arrival in } (t, t + \Delta)) = 1 - e^{-\lambda\Delta} \approx \lambda\Delta$$

So in each small interval Δ is an arrival with probability $\lambda\Delta$!

- Dividing $(0, t)$ into many small intervals of length Δ , the number of arrivals in $(0, t)$ is **binomial** with $n = t/\Delta$ and $p = \lambda\Delta$
- Since n is large and p is small, this number is **Poisson distributed** with parameter $np = \lambda t$ (as Δ tends to 0)

$$P(k \text{ arrivals in } (0, t)) = e^{-\lambda t} \frac{(\lambda t)^k}{k!}, \quad k = 0, 1, 2, \dots$$

- Since density $f(x) = \lambda e^{-\lambda x}$ is maximal for $x = 0$, short inter-arrival times occur more frequently than long ones. So arrivals tend to **cluster**:



- Superposition of many independent rarely occurring arrival flows is (close to) Poisson: this is why **Poisson flows often occur in practice!**
- **Merging of two Poisson flows** with rates λ_1 and λ_2 is again Poisson with rate $\lambda_1 + \lambda_2$, since

$$P(\text{arrival in } (t, t + \Delta)) \approx (\lambda_1 + \lambda_2)\Delta.$$

- **Random splitting of Poisson flows** with rate λ and splitting probability p is again Poisson with rate $p\lambda$, since

$$P(\text{arrival in } (t, t + \Delta)) \approx p\lambda\Delta.$$

Example:

- Two type of jobs arrive at machine for processing, type A and B
- Both job types arrive according to Poisson flows
- Type A jobs arrive at rate 2 jobs per hour, type B with rate 3

Questions:

- What is the probability that during 1 hour no jobs arrive?
- What is the probability that the next job to arrive is type A ?
- What is the probability that during 2 hours at least 2 type B jobs arrive?

- t_d is mean time between **departures from workstation**
- σ_d is standard deviation of time between **departures from workstation**
- Departure rate

$$r_d = \frac{1}{t_d}$$

- Coefficient of variation of time between departures

$$c_d = \frac{\sigma_d}{t_d}$$

- In **serial production lines**:

Departures from workstation i are **arrivals to** workstation $i + 1$, so

$$t_a(i + 1) = t_d(i), \quad c_a(i + 1) = c_d(i)$$

- **Utilization** of workstation with identical m machines

$$u = \frac{r_a t_e}{m}$$

- **Single** machine workstation ($m = 1$)

$$c_d^2 \approx (1 - u^2)c_a^2 + u^2 c_e^2 \quad (\text{weighted average of } c_a^2 \text{ and } c_e^2)$$

- **Multi** machine workstation ($m > 1$)

$$c_d^2 \approx 1 + (1 - u^2)(c_a^2 - 1) + \frac{u^2}{\sqrt{m}}(c_e^2 - 1)$$

Remarks:

- Times between departures are **approximately** independent
- If inter-arrival and process times are **exponential** ($c_a = c_e = 1$), then inter-departure times are independent and exponential ($c_d = 1$)

Question: Does the approximation for c_d make sense?

- If $u \approx 1$ and $m = 1$, then machine nearly always busy, so

$$c_d \approx c_e$$

- If $u \approx 0$, then t_e is very small compared to t_a , so

$$c_d \approx c_a$$

- If $c_a = c_e = 1$, then $c_d = 1$ (as it should)

- Building blocks for describing effects of variability in production lines:
 - Process time variability
 - Flow variability
- **Question:** How to evaluate impact of variability on key performance?
 - WIP
 - cycle time
 - throughput
- **Note:** Process time is often small part of cycle time, extra time is **waiting**
- Fundamental issue is to **understand the underlying causes of waiting**
- Science of waiting: **Queueing theory**
- **Queueing system** consists of three components:
 - arrival process
 - service (production) process
 - queue (buffer)

Components:

- Generator G sends jobs to machine M ;
- Machine M processes these jobs and sends finished jobs to exit E ;
- Exit E is doing some book keeping.

Parameters:

- Mean inter-arrival time t_a of the generator G ;
- Mean process time t_e of the machine M .

Question: What is the throughput δ ?

Note: This model can also be seen as **two-machine zero-buffer mode**

Object type job

14/32

```
type job = int;
```

Jobs are numbered.

```
proc G(chan! job a; real ta):  
  job x;  
  
  while true:  
    a!x;  
    delay ta;  
    x = x + 1;  
  end  
end
```

G generates jobs with constant inter-arrival times t_a .

```
proc M(chan? job a; chan! job b; real te):  
  job x;  
  
  while true:  
    a?x;  
    delay te;  
    b!x;  
  end  
end
```

M processes jobs with constant process times t_e .


```
proc real E(chan? job a; int n):  
  job x;  
  
  while x < n:  
    a?x;  
  end;  
  exit x / time  
end
```

Exit *E* computes throughput over first n jobs.

```
model real GME(real ta, te; int n):  
    chan job a, b;  
  
    run G(a, ta), M(a, b, te), E(b, n)  
end
```

Question: Why is δ for random process times smaller than for constant times?

Let A be inter-arrival time, B the (effective) process time and C cycle time,

$$C = \max\{A, B\}.$$

Then

$$\delta = \frac{1}{E(C)} \leq \frac{1}{\max\{E(A), E(B)\}}$$

with equality only for **constant** A and B .

Randomness leads to **starvation** ($A > B$) and **blocking** ($A < B$)!

Question: How to deal with variations in inter-arrival times and process times?

Answer: Use buffers!

If buffer is sufficiently large, then

$$\delta = \frac{1}{\max\{E(A), E(B)\}}.$$

```
proc B(chan? job a; chan! job b; int N):  
  list job xs;  
  job x;  
  while true:  
    select  
      size(xs) < N, a?x:  
        xs = xs + [x]  
    alt  
      size(xs) > 0, b!xs[0]:  
        xs = xs[1:]  
    end  
  end  
end
```

```
model real GBME(real ta, te; int n, N):  
  chan job a, b;  
  run G(a, ta), B(a, b, N), M(b, c, te), E(c, n)  
end
```

Question: How to calculate throughput in case of buffers?

Answer: Use queueing theory!

- Inter-arrival time is exponential with rate λ
- Process time is exponential with rate μ
- Maximal n jobs in system (including one in G and M).

p_i is long-run probability (or fraction of time) of finding i jobs in the system

Question: How to determine these probabilities?

Answer: Through balance equations!

State i means a total of i jobs in the system, 1 in process and $i - 1$ waiting

Note: One of the waiting jobs in state n occupies machine G (which is blocked)

Balance equations

Flow from state i to $i + 1$ = Flow from state $i + 1$ to i

This yields

$$p_i \lambda = p_{i+1} \mu, \quad i = 0, \dots, n - 1,$$

Solution

$$p_i = p_0 \left(\frac{\lambda}{\mu} \right)^i, \quad i = 0, \dots, n$$

and p_0 follows from normalization

$$1 = \sum_{i=0}^n p_i = p_0 \frac{1 - (\lambda/\mu)^{n+1}}{1 - \lambda/\mu}$$

so

$$p_0 = \frac{1 - \lambda/\mu}{1 - (\lambda/\mu)^{n+1}}.$$

The machine utilization is $u = 1 - p_0$ and the throughput

$$\delta = u\mu = (1 - p_0)\mu$$

Now let buffer capacity n tend to infinity

$$p_0 = 1 - \frac{\lambda}{\mu}$$

and for all $i = 0, 1, 2, \dots$

$$p_i = \left(1 - \frac{\lambda}{\mu}\right) \left(\frac{\lambda}{\mu}\right)^i.$$

Number in system has **geometric distribution** with parameter λ/μ !

Remarks:

- We have to require

$$\lambda < \mu$$

otherwise the system will explode

- **Interpretation:** p_i is long-run fraction of time of finding i jobs in the system

- Throughput

$$\delta = \lambda$$

- Utilization

$$u = 1 - p_0 = \frac{\lambda}{\mu}$$

- Mean Work-In-Process (WIP) level

$$w = \sum_{i=0}^{\infty} i p_i = \frac{\lambda/\mu}{1 - \lambda/\mu} = \frac{u}{1 - u}$$

- Mean Flow time

$$\varphi = \frac{w}{\lambda} = \frac{1/\mu}{1 - \lambda/\mu} = \frac{1/\mu}{1 - u}$$

- Question:** What happens to w and φ as $u \uparrow 1$?

```
proc B(chan? job a; chan! job b):  
  list job xs;  
  job x;  
  while true:  
    select  
      a?x:  
        xs = xs + [x]  
    alt  
      size(xs) > 0, b!xs[0]:  
        xs = xs[1:]  
    end  
  end  
end  
model real GIBME(real ta, te; int n):  
  chan job a, b;  
  run G(a, exponential(ta)), B(a, b),  
    M(b, c, exponential(te)), E(c, n)  
end
```

- Inter-arrival time is **exponential** with mean $t_a = 1/\lambda$
- Process time is **exponential** with mean $t_e = 1/\mu$
- Infinite capacity buffer
- Single machine
- Processing in order of arrival

Then

$$\varphi = \left(\frac{u}{1-u} + 1 \right) \cdot t_e$$

with

- $u = t_e/t_a = \frac{\lambda}{\mu}$

- Inter-arrival time is **exponential** with mean $t_a = 1/\lambda$
- Process time is **general** with mean t_e and standard deviation σ_e
- Infinite capacity buffer
- Single machine
- Processing in order of arrival

Then

$$\varphi = \left(\gamma \cdot \frac{u}{1-u} + 1 \right) \cdot t_e$$

with

- $u = t_e/t_a = \lambda t_e$
- $\gamma = \frac{1}{2} \cdot (1 + c_e^2)$
- $c_e = \sigma_e/t_e$
- Note that $c_e = 1$ for exponential process times!

- Inter-arrival time is **general** with mean t_a and standard deviation σ_a
- Process time is **general** with mean t_e and standard deviation σ_e
- Infinite capacity buffer
- Single machine
- Processing in order of arrival

Then

$$\varphi \approx \left(\gamma \cdot \frac{u}{1-u} + 1 \right) \cdot t_e$$

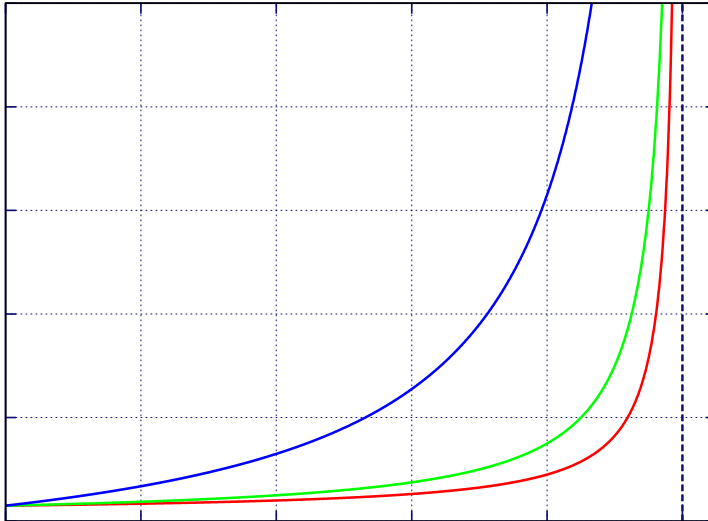
with

- $u = t_e/t_a$
- $\gamma = \frac{1}{2} \cdot (c_a^2 + c_e^2)$
- $c_a = \sigma_a/t_a$
- $c_e = \sigma_e/t_e$

$$\varphi \approx \left(\gamma \cdot \frac{u}{1-u} + 1 \right) \cdot t_e \quad (1)$$

Lessons:

- As u tends to 1 then φ tends to ∞
- As u tends to 1, then approximation (1) becomes **exact**: **relative error** tends to 0;
- As u tends to 1, then *distribution* of flow time becomes **exponential**
- **Summary**:
As system operates close to its maximum capacity, flow times are long and **exponential** with mean (1)
- **Insensitivity**: Mean flow time only depends on mean and standard deviation of inter-arrival times and process times!



Mean flow time φ as function of utilization u , for $0 < u < 1$, and

$c_a = 1$, $c_s = 0, 1, 3$. **Who is who?**