Quantifying Performance Models Part 2

Dr. John Burns

Dept. of Computing
ITT Dublin

November 6, 2017

Motivating Problem I

Conceptually, Little's Law is quite simple and intuitively appealing. We describe the result by way of an analogy:

Little's Law Analogy

Consider a coffee shop. Customers arrive at the coffee shop, stay for a while, and leave. Little's result states that the average number of folks in the coffee shop is equal to the departure rate of customers from the coffee shop times the average time each customer stays in the coffee shop. See Fig. 1.

Little's Law I

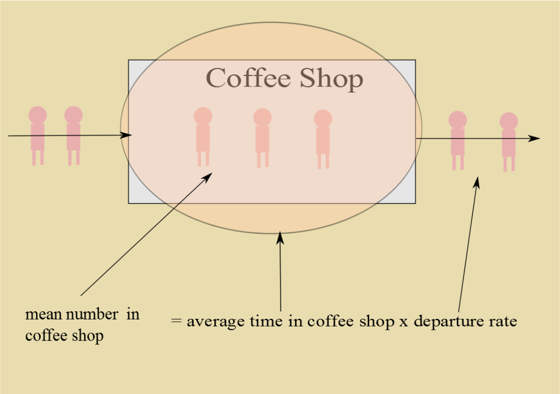

Figure 1: Little's Law $N = XR$

Little's Law

- This result applies across a wide range of assumptions. For instance, consider a deterministic situation where a new customer walks into the coffee shop every hour on the hour.
- Upon entering the coffee shop, suppose that there are three other customers in the coffee shop.
- Suppose that the barista regularly kicks out the customer who has been there the longest, every hour at the half hour.
- Thus, a new customer will enter at **9:00, 10:00, 11:00, ...,** and the oldest remaining customer will be booted out at **:** **9:30, 10:30, 11:30, ... ,**

Little's Law

It is clear that the average number of persons in the coffee shop will be $3\frac{1}{2}$. Thus, via Little's Law, *the average number in the coffee shop* = *departure rate times the average time spent in the coffee shop*: $3\frac{1}{2} = 1 \times 3\frac{1}{2}$

- In fact, Little's Law holds as long as customers are not destroyed or created.
- For example, if there is a fight in the coffee shop and someone gets killed or a if a pregnant woman goes into the coffee shop and gives birth, Little's Law does not hold.
- Little's Law applies to any "black box", which may contain an arbitrary set of components.
- If the box contains a single resource (e.g., a single CPU, a single coffee shop) or if the box contains a complex system (e.g., the Internet, a city full of coffee shops and grocery shops), Little's Law holds.

Little's Law

Thus, Little's Law can be restated as

Definition

Little's Law The average number of jobs inside a system is equal to the departure rate times the average time spent in the system.
$N = XR$

Little's Law Restated

Example

Question: Consider the database server in the early example, and assume that during the same measurement interval the average number of database transactions in execution was 16. What was the response time of database transactions during that measurement interval? **Answer:** The throughput of the database server was already determined as being 3.8 tps. Apply Little's Law and consider the entire database server as the box. The average number in the box is the average number N of concurrent database transactions in execution (i.e., 16). The average time in the box is the average response time R desired. Thus, $R = N/X_0 = 16/3.8 = 4.2$ sec.

Interactive Response Time Law I

- Consider an interactive system composed of M clients each sitting at their own workstation and interactively accessing a common database server system.
- Clients work independently and alternate between "thinking" (i.e., composing requests for the server) and waiting for a response from the server.
- The average think time is denoted by Z and the average response time is R. See Fig. 2.
- The think time is defined as the time elapsed since a customer receives a reply to a request until a subsequent request is submitted.
- The response time is the time elapsed between successive think times by a client.

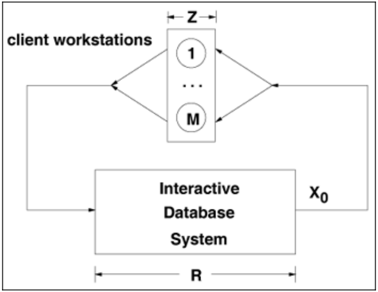

Figure 2: Interactive computer system.

Interactive Response Time Law I

- Let $\bar{M}$ and $\bar{N}$ be the average number of clients thinking and the average number of clients waiting for a response, respectively.
- By viewing clients as moving between workstations and the database server, depending upon whether or not they are in the think state, $\bar{M}$ and $\bar{N}$ and represent the average number of clients at the workstations and at the database server, respectively.
- Clearly, $M = \bar{M} + \bar{N}$ since a client is either in the think state or waiting for a reply to a submitted request.

Interactive Response Time Law II

- By applying Little's Law to the box containing just the workstations,

$$\bar{M} = X_0 Z \tag{1}$$

since the average number of requests submitted per unit time (throughput of the set of clients) must equal the number of completed requests per unit time (system throughput X0). Similarly, by applying Little's Law to the box containing just the database server,

$$\bar{N} = X_0 R \tag{2}$$

Interactive Response Time Law III

where R is the average response time. By adding Eqs (2) and (1) we get:

$$\bar{M} + \bar{N} = M = X_0(Z + R) \tag{3}$$

or equivalently:

$$R = \frac{M}{X_0} - Z \tag{4}$$

This is an important formula known as the **Interactive Response Time Law**.

Interactive Response Time Law: Example I

Example 1

If 7,200 requests are processed during one hour by an interactive computer system with 40 clients and an average think time of 15 sec, the average response time is: $R = \frac{40}{7200/3200} - 15 = 5$

Interactive Response Time Law: Example II

Example 2

A client/server system is monitored for one hour. During this time, the utilization of a certain disk is measured to be 50%. Each request makes an average of two accesses to this disk, which has an average service time equal to 25 msec. Considering that there are 150 clients and that the average think time is 10 sec, what is the average response time? The known quantities are: $U_{disk} = 0.5$, $V_{disk} = 2$, $S_{disk} = 0.025$ sec, $M = 150$, and $Z = 10$ sec. From the Utilization Law, $U_{disk} = S_{disk} X_{disk}$ Thus, $X_{disk} = 0.5/0.025 = 20$ requests/sec. From the Forced Flow Law, $X_0 = \frac{X_{disk}}{V_{disk}} = \frac{20}{2} = 10 req/sec$ Finally, from the Interactive Response Time Law, $R = \frac{M}{X_0} - Z = \frac{150}{10} - 10 = 5 sec$

Summary of Operational Laws I

The following laws summarise the material covered in this section:

Utilisation:

$$U_i = S_i \times X_i = \lambda_i \times S_i \tag{5}$$

Forced Flow:

$$X_i = V_i \times X_0 \tag{6}$$

Service Demand:

$$D_i = V_i \times S_i = \frac{U_i}{X_0} \tag{7}$$

Little's Law

$$N = XR \tag{8}$$

Interactive Response Time Law

$$R = \frac{M}{X_0} - Z \tag{9}$$