Computer Systems Performance Analysis and Benchmarking (37-235)

Analytic Modeling

Simulation

Measurements / Benchmarking

Lecture by:

Prof. Thomas Stricker

Assignments/Projects:

Christian Kurmann

Textbook:

Raj Jain, "The Art of Computer Systems Performance Analysis", 1991 Wiley & Sons, New York

Topic of Today:

- **Transient Removal**
- **Monitors**
- **Random Generators for Simulation**

Monitors

Used by:

- systems programmer (optimization)
- systems manager (find bottleneck)
- systems manager (tune system)
- systems analyst (characterize work)
- systems analyst (find model params.)

Terms:

- Event = change in system state
- Trace = log of events incl. time
- Overhead = perturbation by monitor
- Domain = activities observed
- Input rate = max frequency of event
- Resolution = coarseness of information
- Input width = number of bit sampled

Monitor Classifications

1. Dimension

- Event driven
- Sampling
- Hybrid

2. Dimension

- On-line
- Batch

Methods:

- Software Monitors
- Hardware Monitors

Design Issues (Software)

- Activation Mechanism - On/Off switch
- Trap instruction - Timer Interrupt
- Trace mode - Priority
- Buffer size - overflow - number of bufs
- Data compression/analysis on the fly
- Language
- Monitoring abnormal events

Design Issues (Hardware)

- Probes
- Counters - Timers
- Logic Elements
- Comparators
- Mapping Hardware
- Tape/Disks

Software vs. Hardware

Criterion	Hardware Monitor	Software Monitor
Domain	Difficult to monitor operating system events.	Difficult to monitor hardware events unless recognizable by an instruction.
Input rate	Sampling rates of 10^5 per second possible.	Sampling rate limited by the processor MIPS and overhead required.
Time resolution	10 nanoseconds is possible.	Generally 10 to 16 milliseconds.
Expertise	Requires intimate knowledge of hardware.	Requires intimate knowledge of software.
Recording capacity	Limited by memory and secondary storage. Not a problem currently.	Limited by overhead desired.
Input width	Can record several simultaneous events.	Cannot record several simultaneous events unless there are multiple processors.
Monitor overhead	None	Overhead depends upon the input rate and input width. Less than 5% adequate and more than 100% possible.
Portability	Generally portable.	Specific to an operating system.
Availability	Monitoring continues even during system malfunction or failure.	Cannot monitor during system crash.
Errors	Possible to connect the probes to wrong points.	Once debugged, errors are rare.
Cost	High	Medium

Distributed Systems Monitor

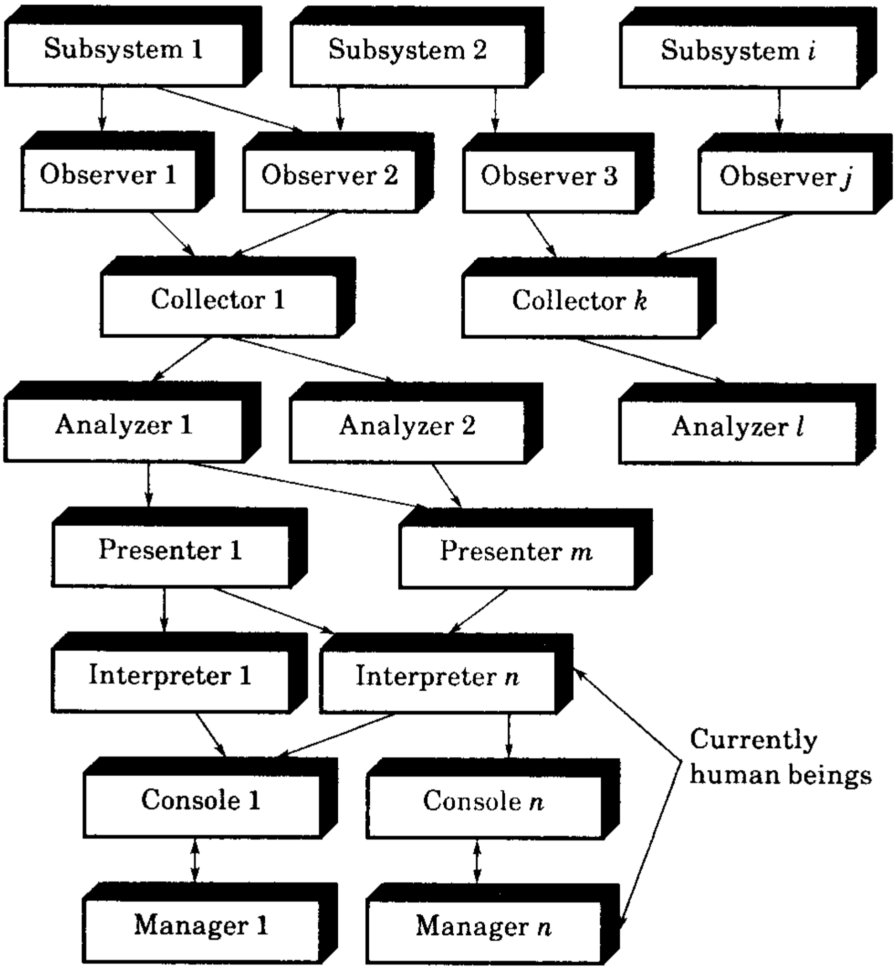


FIGURE 7.2 Components of a distributed-system monitor.

Random Generators

- Used for simulators
- True randomness is not desired (reproducability for debugging)
- Statistical properties must be random
- No cryptographic strength required

Example

$$x_n = f(x_{n-1}, x_{n-2}, \ldots)$$

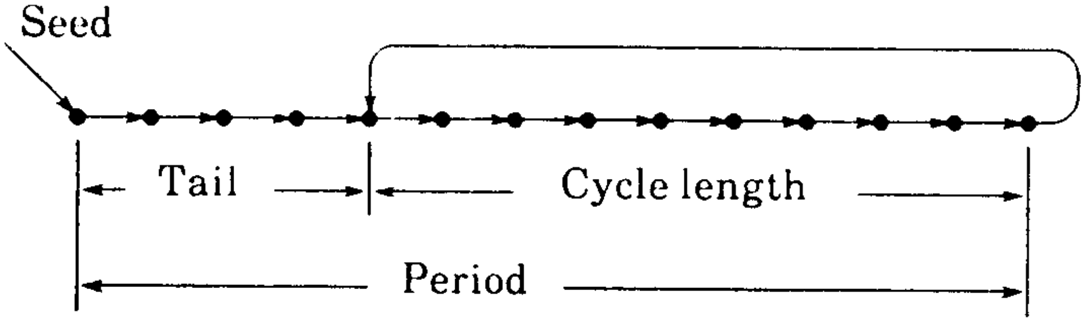
One such function is

$$x_n = 5x_{n-1} + 1 \bmod 16$$

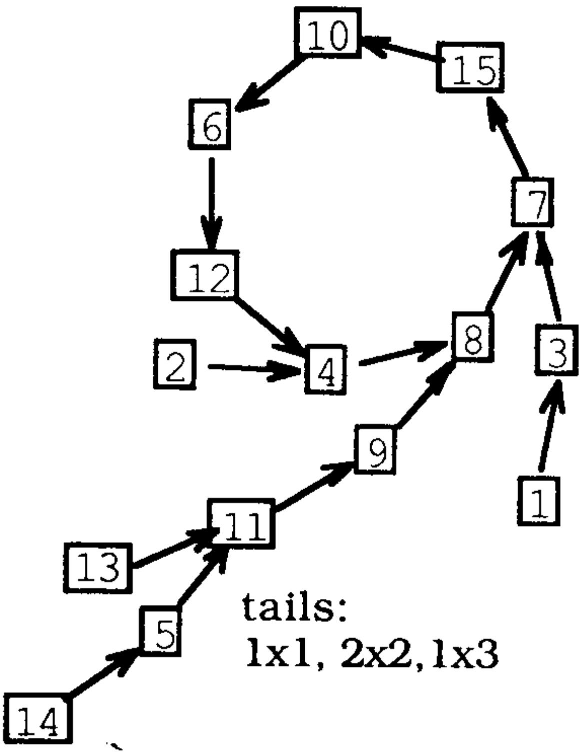
$$x_1 = 5(5) + 1 \bmod 16 = 26 \bmod 16 = 10$$

10, 3, 0, 1, 6, 15, 12, 13, 2, 11, 8, 9, 14, 7,
4, 5, 10, 3, 0, 1, 6, 15, 12, 13, 2, 11, 8, 9,

cycle length **seed.**



Cycles



Properties

1. *It should be efficiently computable.* Since simulations typically require several thousand random numbers in each run, the processor time required to generate these numbers should be small.
2. *The period should be large.* A small period may cause the random-number sequence to recycle, resulting in a repeated event sequence. This may limit the useful length of simulation runs.
3. *The successive values should be independent and uniformly distributed.* The correlation between successive numbers should be small. Correlation, if significant, indicates dependence.

- Linear-congruential generators
- Tausworthe generators
- Extended Fibonacci generators
- Combined generators

Linerar Congruence Generators

$$x_n = a^n \bmod m$$

$$x_n = a x_{n-1} \bmod m$$

$$x_n = a x_{n-1} + b \bmod m$$

$$a = 23 \text{ and } m = 10^8 + 1.$$

full-period generator.

lower autocorrelation

$$x_n = (2^{34} + 1)x_{n-1} + 1 \bmod 2^{35}$$

$$x_n = (2^{18} + 1)x_{n-1} + 1 \bmod 2^{35}$$

Multiplicative LCG

$$x_n = a x_{n-1} \bmod m$$

$$m = 2^k$$

$$x_n = 5x_{n-1} \bmod 2^5$$

$$x_n = 7x_{n-1} \bmod 2^5$$

Examples

$$m \neq 2^k$$

$$x_n = 3x_{n-1} \bmod 31$$

$$5^3 \bmod 31 = 125 \bmod 31 = 1$$

$$x_n = 7^5 x_{n-1} \bmod (2^{31} - 1)$$

Shiftregister Generators

TAUSWORTHE GENERATORS

$$b_n = c_{q-1}b_{n-1} \oplus c_{q-2}b_{n-2} \oplus c_{q-3}b_{n-3} \oplus \cdots \oplus c_0 b_{n-q}$$

$$x^7 + x^3 + 1$$

$$D^7 b(n) + D^3 b(n) + b(n) = 0 \bmod 2$$

$$b_{n+7} + b_{n+3} + b_n = 0 \bmod 2,$$

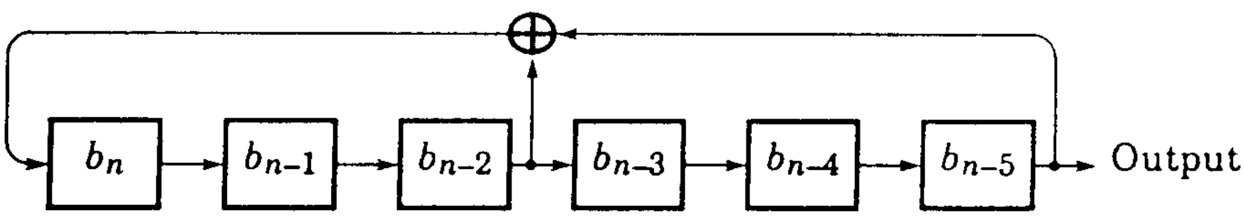
$$b_7 = b_3 \oplus b_0 = 1 \oplus 1 = 0$$

$$b_8 = b_4 \oplus b_1 = 1 \oplus 1 = 0$$

$$b_9 = b_5 \oplus b_2 = 1 \oplus 1 = 0$$

$$b_{10} = b_6 \oplus b_3 = 1 \oplus 1 = 0$$

$$b_{11} = b_7 \oplus b_4 = 0 \oplus 1 = 1$$



1111111 0000111 0111100 1011001 0010000
0010001 0011000 1011101 0110110
0000110 0110101 0011100 1111011
0100001 0101011 1110100 1010001
1011100 0111111 1000011 1000000.

period of 127.

Chi-Square Test

$$D = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i}$$

identically distributed (**IID**) $U(0, 1)$,

testing random-variate generators.

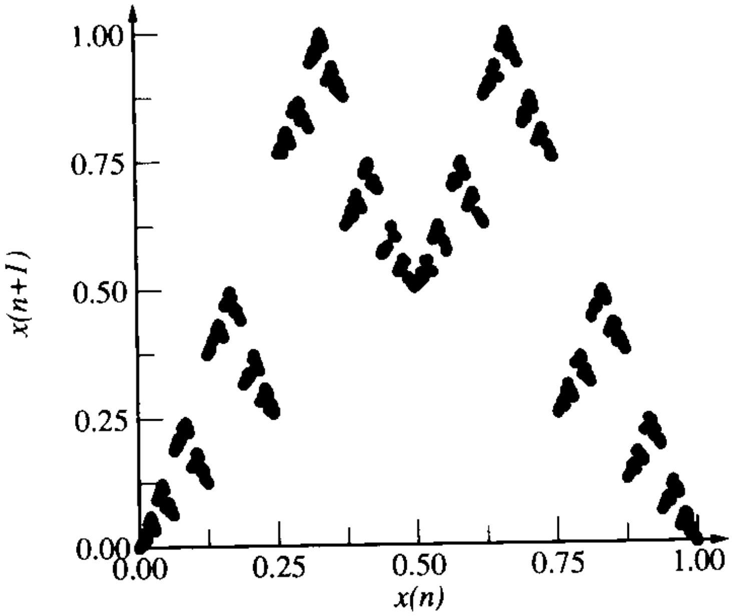
D has a chi-square distribution

level of significance α

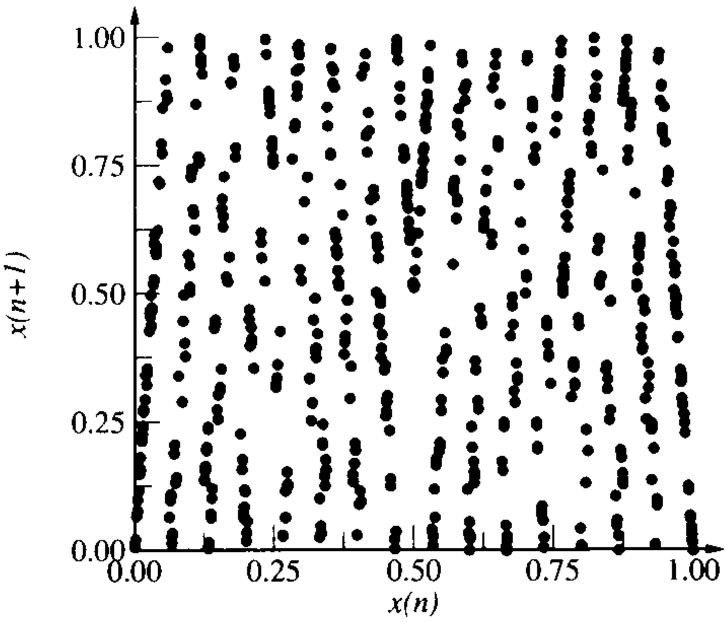
computed D is *less* than the $\chi^2_{[1-\alpha;k-1]}$

read from Table A.5

Graphical Tests

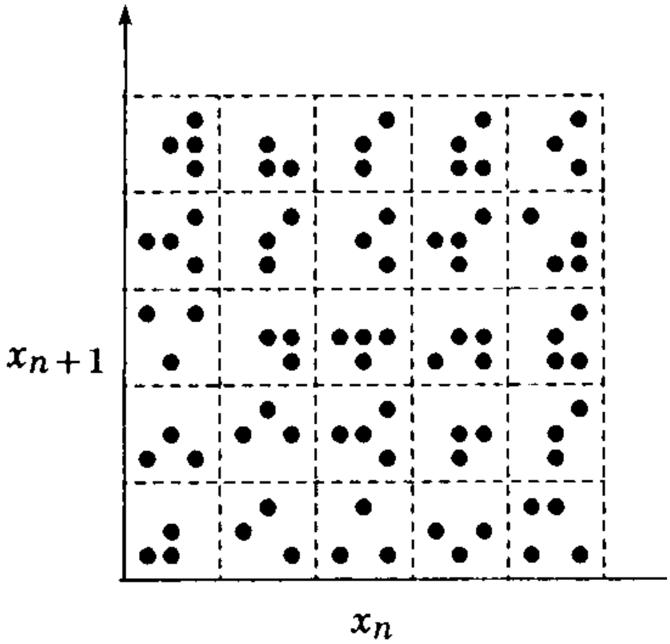


Plot of overlapping pairs
Tausworthe generator $x^{15} + x + 1$.



Plot of overlapping pairs
Tausworthe generator $x^{15} + x^4 + 1$.

Graphical Tests



Two-dimensional uniformity