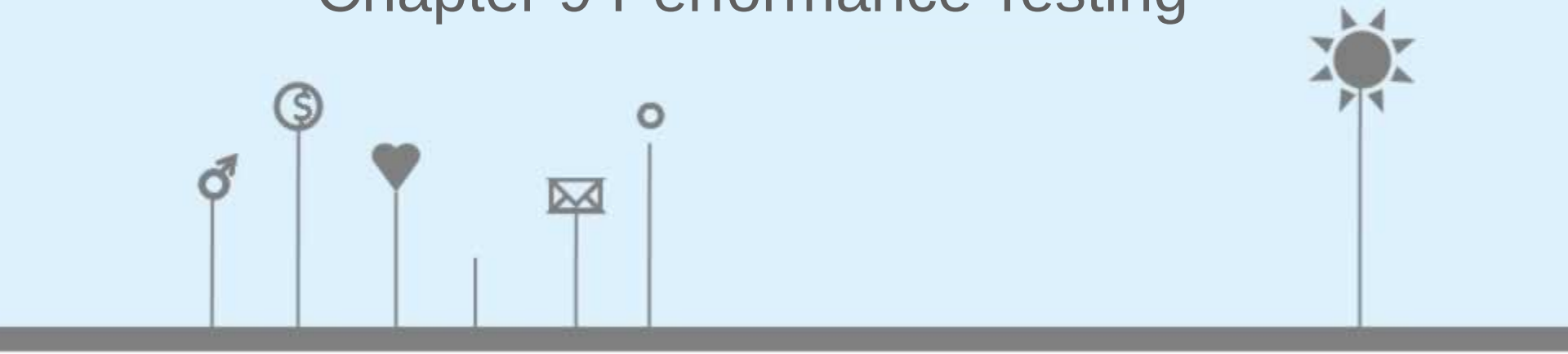
Software College Northeastern University

Software Quality Assurance and Testing

Chapter 9 Performance Testing



wuchenni@qq.com

Contents

Chapter 9

Performance Testing

9.1 What is performance?

销售甲：*“我们的系统功能那是相当的强大，完全能够满足您所有的功能需求。”*

客户乙（使用后……）：*“果然不错！合作愉快！”*

……（一个星期后）

项目经理甲：*“上星期不是试过了吗？怎么会慢呢？”*

客户乙：*“1个人用是没问题，但是我们公司有1000个员工使用这个系统，一个查询的操作要等1分钟，还时不时就崩溃，简直无法忍受！！”*

销售甲：……

9.1 What is performance?

- Performance(性能) is the degree to which a system or component accomplishes its designated functions within given constraints, such as speed, accuracy, or memory usage.
- Performance is an external quality based upon user requirements and the user's view of the operational system.
- Performance is also especially critical for real-time systems in which actions must complete within a specified time limit for correct operation.

9.1 What is performance?

- Commonly used performance indicators
  - Example：Time efficiency, space efficiency, transaction operating performance, I/O operating performance, the database performance, memory performance, the initialization/exit time and resources utilization, etc.
  - Introduce some common performance indicators : Latency, Transaction processing time, Maximum transaction processing time, Transaction operating time, Database performance, The largest consumption of memory, Peak time of memory, Resource consumption.

9.1 What is performance?

- **Latency:**
  - The time interval between the instant at which an instruction control unit issues a call for data and the instant at which the transfer of data has started; the delay between request and completion of an operation.
- **Transaction processing time:**
  - The transaction processing time refers to the running time needed to complete a transaction, used to evaluate the transaction processing efficiency. Usually, the shorter time a transaction needed, the higher the efficiency is.

9.1 What is performance?

- **Maximum transaction processing time:**
  - It is a very important performance indicators, we first need analyze which transactions spend more time, then tested out the time spent, it takes the most time is the biggest transaction processing time.
- **Transaction operating time:**
  - Mainly used to evaluate operating time require for the user to operate the transaction processing, It mainly reflect the efficiency of the user operation. Test the transactions operating time, need a timer, test different users spend time, finally take average.

9.1 What is performance?

- **Database performance :**
  - Database performance indicators generally include the time needed for search, insert, and delete and update the database, etc, due to database performance is often the bottleneck of software system performance, so it is very necessary to test database performance.
- **The largest consumption of memory:**
  - The largest consumption of memory shall be equipped what level of hardware to run software, it is the direct reflection of hardware cost. The largest consumption of memory can be tested with a performance monitor.

9.1 What is performance?

- **Resource consumption:**
  - The amount of memory or disk space consumed by the application.
- **Peak time of memory:**
  - The running time for software consumes in peak period of memory. If it is closer for the software used in the peak of memory and system total memory, the efficiency of software will be greatly reduced. So, reduce the peak time of memory can improve software performance.

9.1 What is performance?

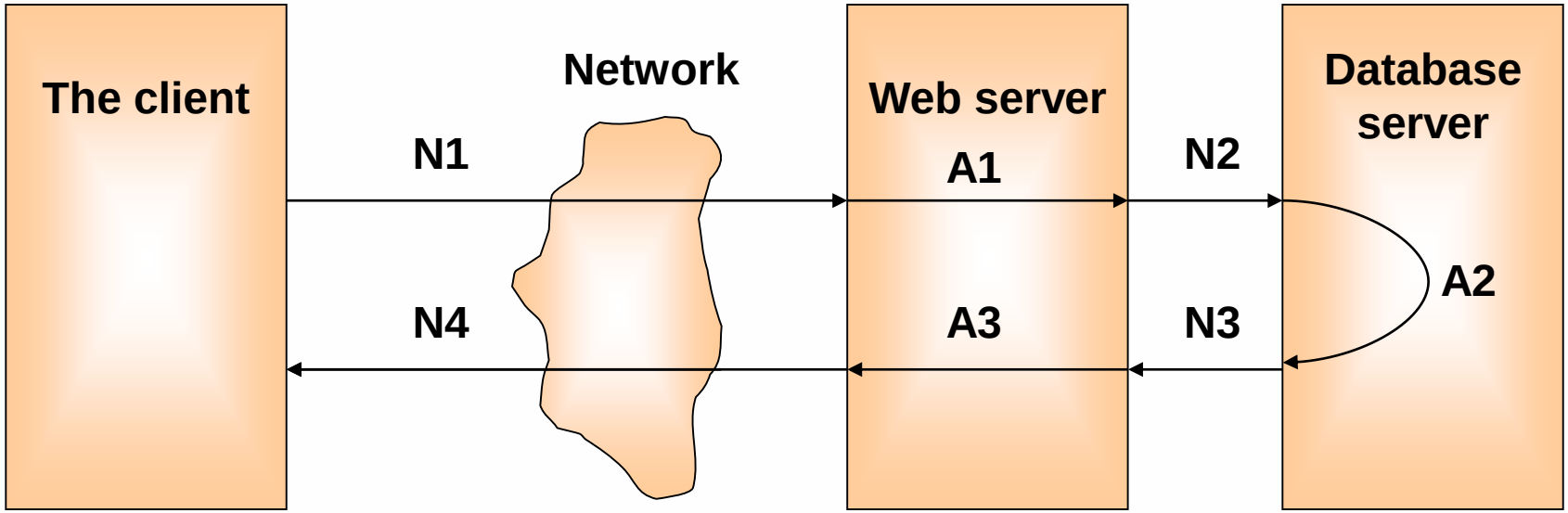
- Now web applications are very extensive, introduces several indicators below all about web applications performance, such as: Concurrent users, Response time, Throughput and Dormancy (休眠)time, etc.

9.1 What is performance?

- **Response time:**
  - Response time refers to the time that needed for requests make a response, it can divide into transmission time, application latency time, database latency time and present time, etc.
- **System response time:**
  - Application system consumptive time from sending a request to client receiving data.
- **Present time:**
  - Present time depends on the consumption time for presenting page after the data received by the client.

9.1 What is performance?

- Page response time decomposition for a web application



- Page response time divide into transmission time (N1+N2+N3+N4) and application latency time (A1+A2+A3).
- Application latency time divide into database latency time(A2) and Web server latency time(A1+A3).

9.1 What is performance?

- 一个用户访问一个独立页面时，页面给予用户响应并有所显示时，等待不同时间的态度分布：
  - 10秒钟之内，84%的用户愿意等待
  - 15秒钟之内，51%的用户愿意等待
  - 20秒钟之内，26%的用户愿意等待
  - 30秒钟之内，5%的用户愿意等待

9.1 What is performance?

- Generally speaking, response time of the website have a few standards like 2s, 5s and 10s.
- 2s — Very attractive

- 5s — Relatively good

- 10s — Cannot tolerate

9.1 What is performance?

- Concurrent users
  - Generally divide kind of case:
    - Strictly concurrent, that all users at the same time doing the same thing or operation.
    - Broad category of concurrent, multiple users send out a request to system or make a operation, but these operations can be the same, also can be different.

9.1 What is performance?

- Performance testing is concern to concurrent users of business, how many concurrent users should be set from the angle of business is proper.
- The following are the concurrent users estimate formula:

$$C = \frac{nL}{T}$$

  - C — Concurrent users on average
  - n — The number of logon session
  - L — The average length of logon session
  - T — Inspection period length

$$\hat{C} \approx C + 3\sqrt{C}$$

  - $\hat{C}$ — The peak concurrent users

9.1 What is performance?

- 例题：
- 一个软件系统每天约有400个用户访问。用户在一天之内有8小时内使用该系统，从登录到退出的平均时间为4小时，请计算该系统的并发用户数和并发用户的峰值各是多少?
- 分析：根据公式C=400×4／8=200

$$\hat{C} = 200 + 3\sqrt{200} = 242$$

9.1 What is performance?

- Throughput refers to the process of disposable（一次性）performance testing can test on the network combined amount of data transmission.
- Generally speaking, throughput measure with requests /sec or page numbers/sec.

9.1 What is performance?

- Throughput indicator has the following two roles:
  - (1) Assist to the design performance test scenes to measure performance test scene whether reach to expected design goal.
  - (2) Assist to analysis performance testing bottleneck.

9.1 What is performance?

- Between throughput and concurrent users have certain relationship.
- Computation formula is as follows：

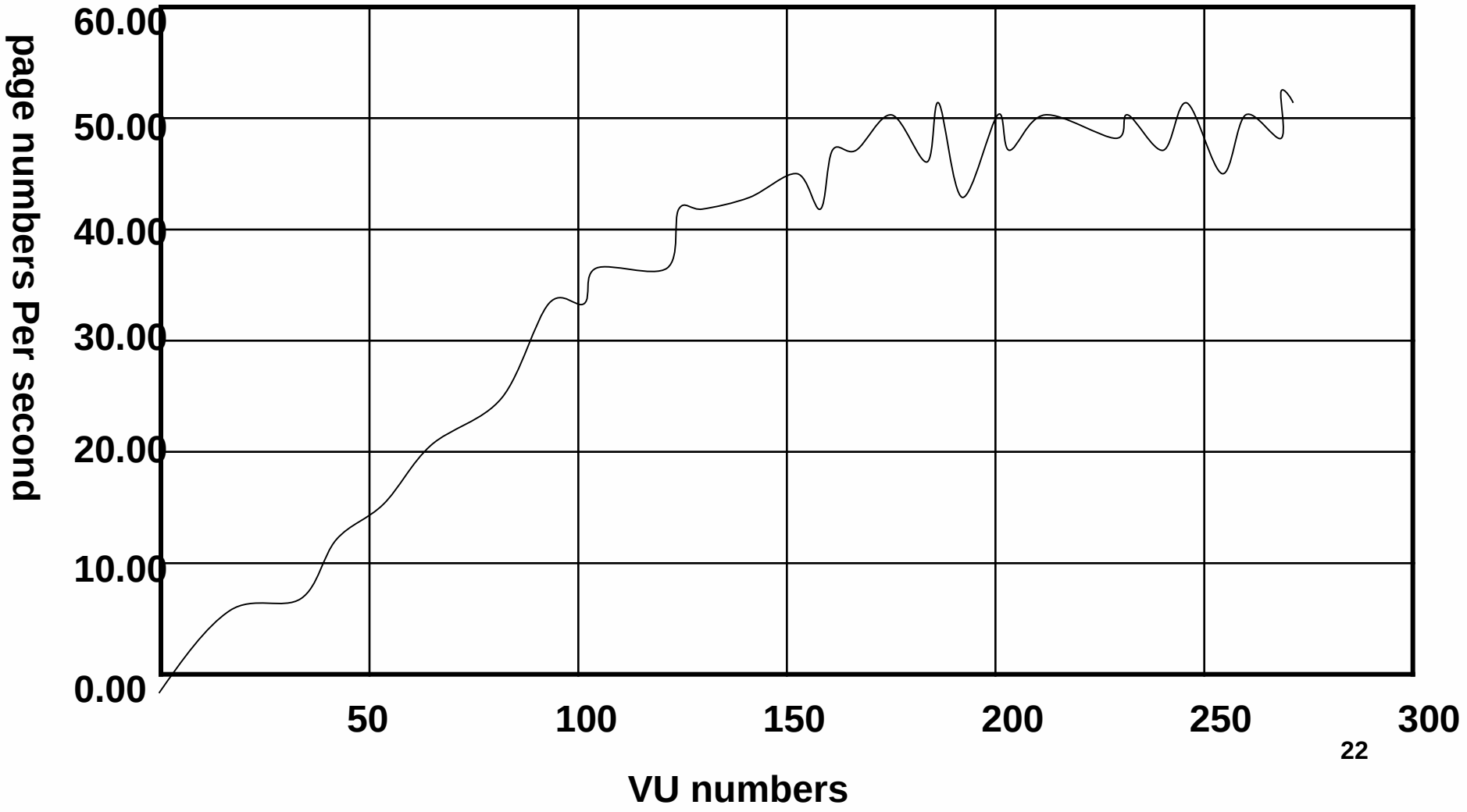
$$F = \frac{N_{vu} \times R}{T}$$

  - F — Throughput
  - N_{vu} — Virtual users number
  - R — Each virtual users send requests
  - T — Time for performance testing

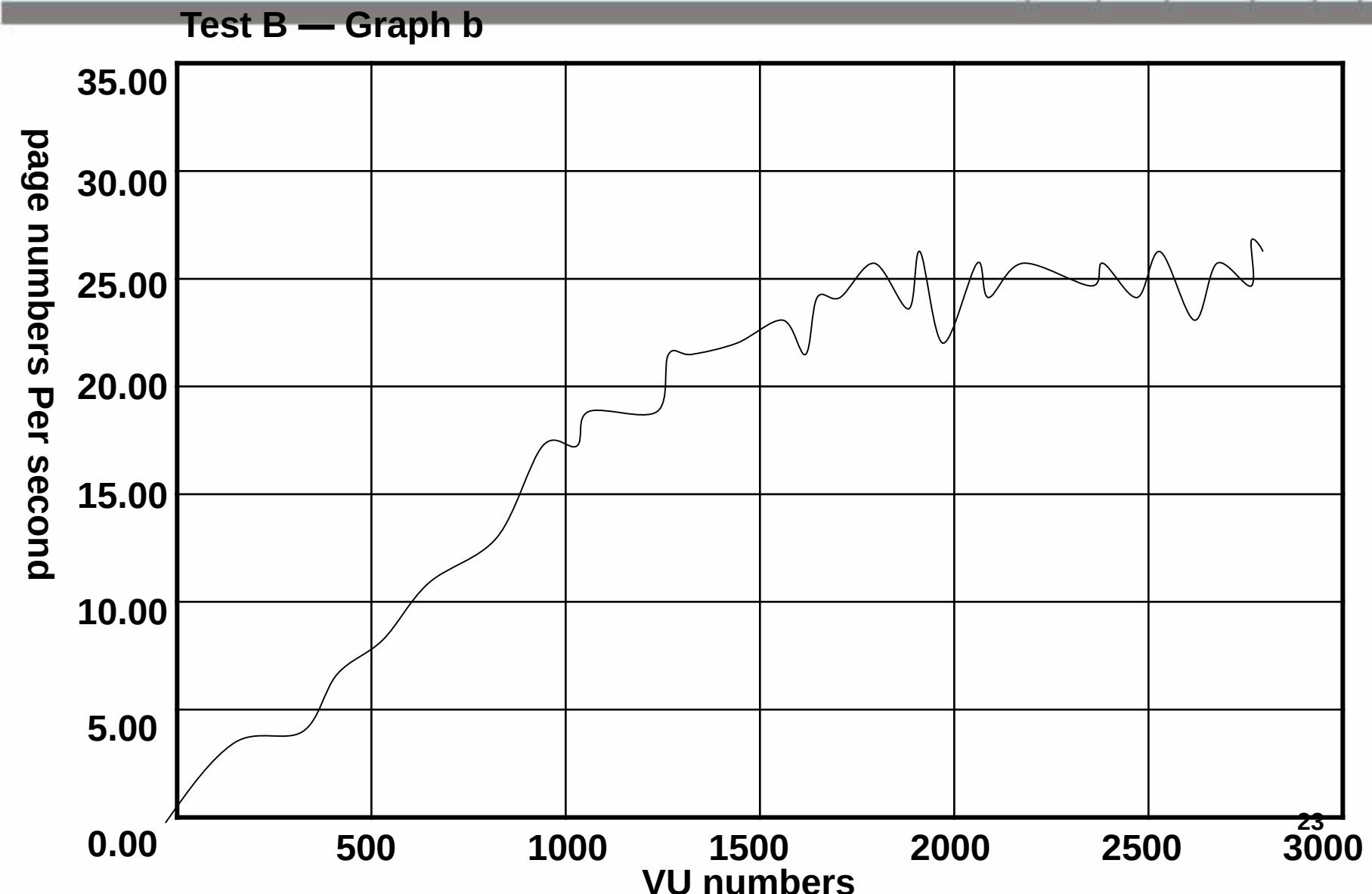
9.1 What is performance?

- Meet performance bottlenecks status, throughput and VU number will not meet formula.
- Below is "throughput - VU number" correlation chart examples.
- A and B indicate two different performance tests for the same software system.
  - Test A have 100 concurrent users, each VU send 1 request pre 1 second;
  - Test B have 1000 concurrent users, each VU send 1 request pre 10 seconds.

9.1 What is performance?



9.1 What is performance?



9.1 What is performance?

- Performance counters and resource utilization
  - Performance counters some of the data indicators is to describe server or operating system performance, with the role of monitoring and analysis.
  - Resource utilization rate refers to the use of various resources.
    - Example:1000 concurrent users access system, web server CPU occupancy rate is 68%, memory occupancy rate is 55%. Where 68% and 55% is resources utilization.
    - Through the resource utilization of lateral comparison（横向比较）, may find performance bottlenecks.

9.1 What is performance?

- Dormancy time
  - Dormancy time also called thinking time, it is point to the user requests time intervals.
  - Therefore, automation test emulates user operation, test scripts must let each operating interval for a period of time, in operating statements setting dormancy time by function Think(), realize waiting time between two operations.
  - There are certain relationship between the dormancy time, iterations, concurrent users and throughput.

9.1 What is performance?

$$F = \frac{N_{vu} \times R}{T} \qquad R = \frac{T}{T_s}$$

- R — Each virtual users send requests
- T_s — Thinking Time
- Throughput and N_{vu} direct proportion, and inversely proportional to T_s.

9.1 What is performance?

- In the actual test, can be calculated thinking time by follows steps:
  - (1) First calculated system concurrent users.
  - (2) Statistics a system average throughput.
  - (3) Statistics a average of each user the request numbers.
  - (4) According to calculation get the thinking time.

9.2 What is performance testing?

- Performance testing(性能测试) is the testing conducted to evaluate the performance of a system with specified performance requirements.
- Performance tests determine the runtime “behavior” of the application and its supporting infrastructure, under certain conditions.
- Performance testing is used to measure several system characteristics, such as processing speed, response time, resource consumption, throughput, and efficiency.

9.2 What is performance testing?

- Performance testing profiles
  - One of the key activities in performance testing is to decide what type of load or loads to place on the system/application.
  - There are many variations of performance testing that we can select from so we first have to determine what is needed.

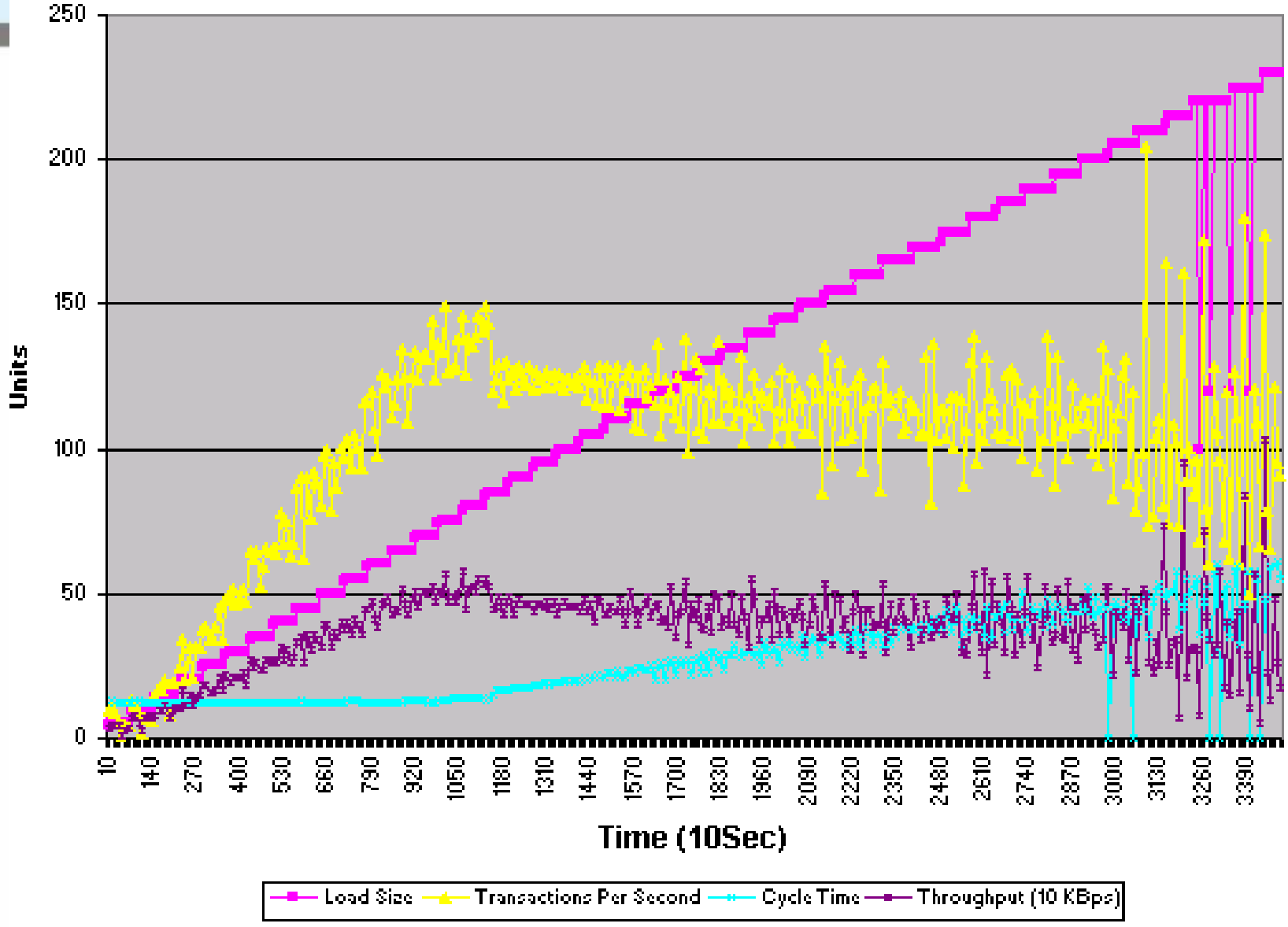
9.2 What is performance testing?

- Load Testing
  - Load Testing determine the applications behavior under load, up to and including its limits (not just at its limits).
  - Load testing specifically refer to the load size (number of concurrent users) and related values.

9.2 What is performance testing?

- Load Testing
  - Load testing is simulated practical software system load conditions under the system load, through the continuous load additions (increasing the number of virtual concurrent users) or other loading additions ways to observe the performance indicators under different load system , such as response time and data throughput, system resources utilization (CPU and memory), to test the behavior and characteristics, discovery possible performance bottlenecks of the system, memory leaks and other problems like cannot achieve synchronization.

Load Test Report



9.2 What is performance testing?

- Load testing loading mode
  - One-time loading
  - Increasing loading
  - High and low mutation loading
  - Random loading method

9.2 What is performance testing?

- **One-time loading:**
  - One-time loading a number of users, at a certain period of time keep on running .
  - For example, in the morning when start working it is time to access web site or land the website very concentrated, basically fall into flat load mode.

9.2 What is performance testing?

- **Increasing loading:**
  - Increase the users by a regular and gradually way to add some new users per second, crisscross rise.
  - Assistant by this kind of loading to test, easy to find inflection points of performance, namely performance bottlenecks.

9.2 What is performance testing?

- **High and low mutation loading**
  - At a certain time the number of users in large quantity, suddenly dropped to be very low，for a long time, and suddenly added to high point, repeated several times.
  - Assistant by this kind of loading to test, easy to find problems like memory leaking and resources releasing.

9.2 What is performance testing?

- **Random loading method**
  - By random algorithm automatically generated within the scope of a certain amount changes, dynamic loading, and this way may be most closely and actual conditions of a loading mode.
  - Although not easy simulation the peak moment while system running, but can simulate the state of system running for a long time.

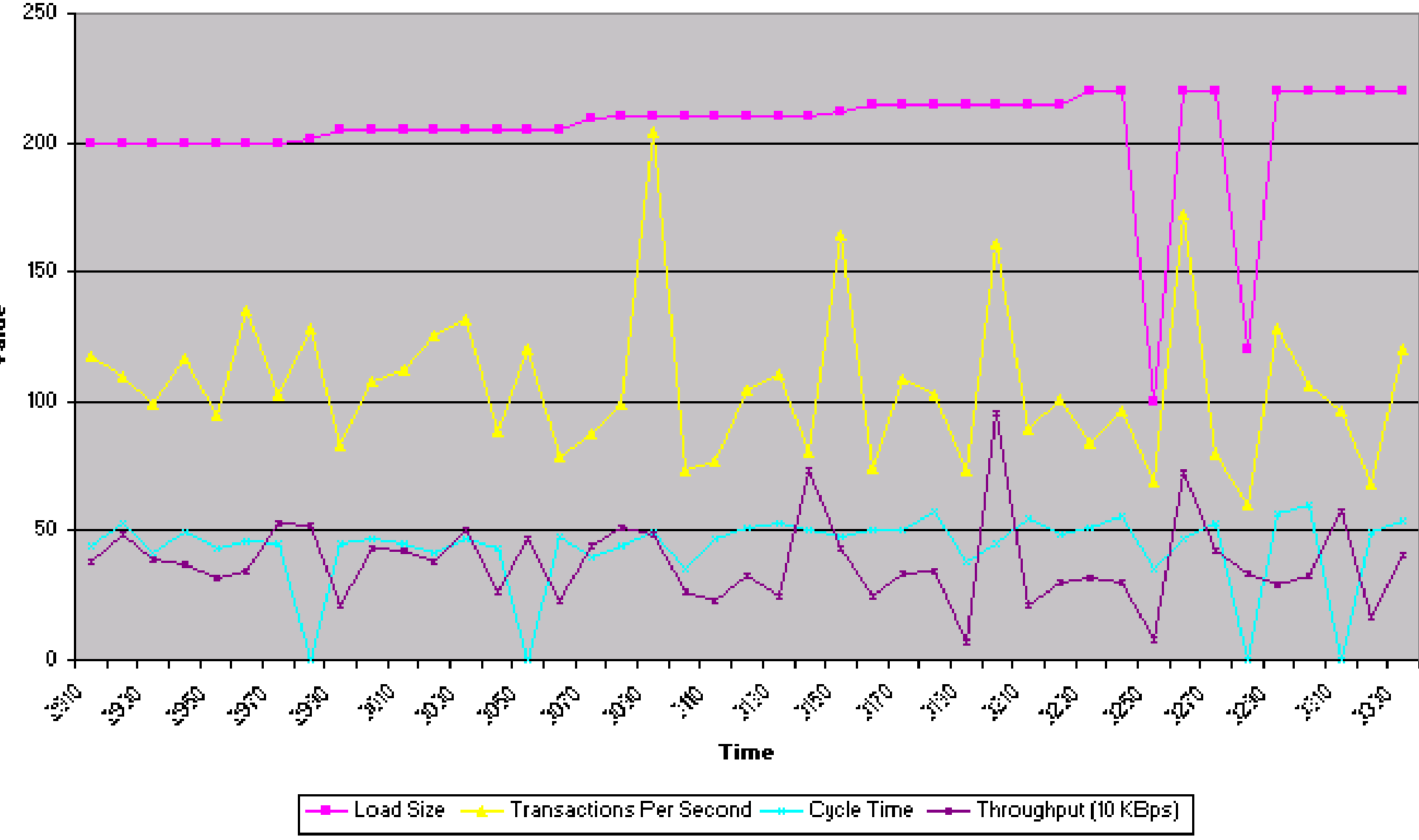
9.2 What is performance testing?

- Stress Testing
  - Stress testing determines the applications ability in strong load test (handle large amounts of data, large numbers of concurrent users).
  - Stress testing can be much more successful with a full load applied to the server.
  - Running addition applications on the server in conjunction with the client side tests is an additional form of stress testing.

9.2 What is performance testing?

- Stress testing can divide into two types:
  - High load long time (such as above 24 hours) the stability stress testing.
  - Limit load cases cause the system collapse destructive stress testing.
- Microsoft testing practice experience shows that if the software products pass the 72 hours stress testing, problems after 72 hours is extremely unlikely. So, 72 hours become Microsoft product stress testing time marks.

Stress Test Report



9.2 What is performance testing?

- Performance, Load, and Stress Testing ?
  - let's apply these terms while testing a car:
    - Performance Testing - a car does 0-60mph in a certain number of seconds under normal conditions.
    - Load Testing – The car can pull a trailer weighing _x_ - tons before it can no longer move.
    - Stress Testing - The car continues to work fine driving 60 mph for two hours, in 120 degree temperatures, with only 1 gallon of coolant in the radiator.

9.3 What are the types of performance tests?

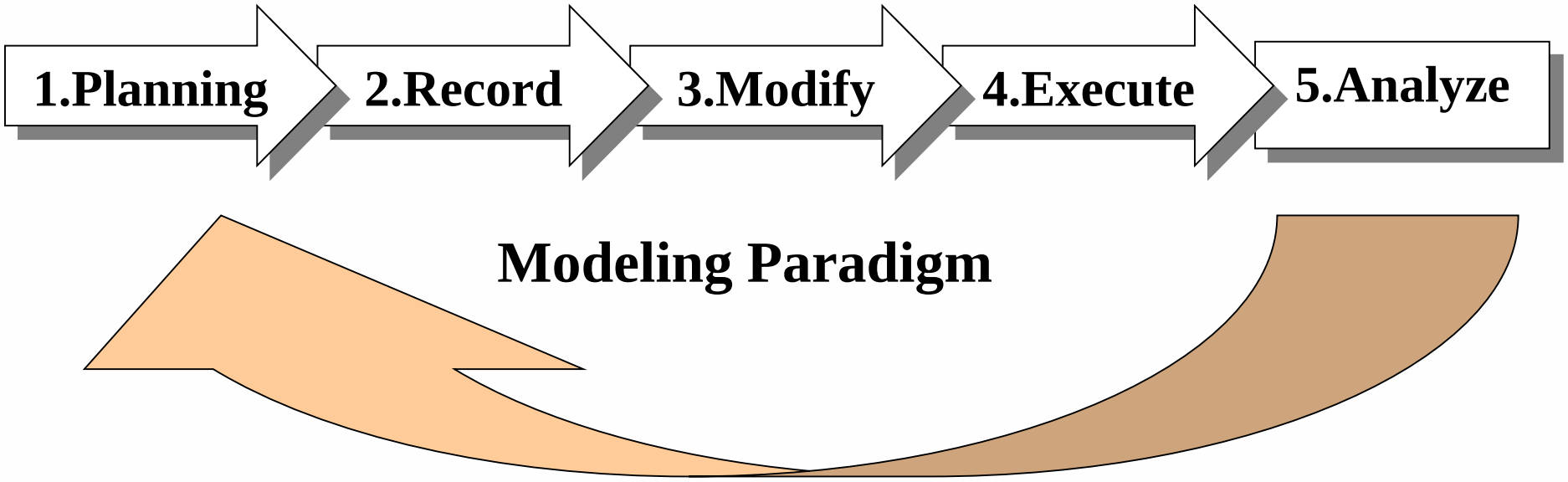
- The three most common types of software performance testing include:
  - Testing which is driven by what we want to measure.
  - Testing which is based on the source or type of the load.
  - Testing to stress the system or find its limits.

9.3 What are the types of performance tests?

- Testing which is driven by what we want to measure
  - Response time testing; Throughput testing
- Testing which is based on the source or type of the load
  - **Load variation/bounce testing**
    - Vary the test load during the performance measurement, to reflect the typical pattern of how the real load ebbs and flows over time.
  - **Calibration(校准) testing**
    - Check a system against mandated requirements such as those from a government regulatory agency.
- Testing to stress the system or find its limits
  - **Deadlock testing**
    - Running volume tests to detect if there are resource contentions in the system such as multiple requestors trying to access the same data base and causing the system to lock up the user request due to conflicts in areas like data base record locks.

9.4 How to conduct performance testing?

- The process of performance testing?



9.4 How to conduct performance testing? Step1:Planning

- Understanding load
  - Once we have an idea of what we want to accomplish
    - The general type of testing we want.
    - What tests we want to run.
  - Now we have to decide on the volume of activity and the composition of that volume, load (i.e. workload)

9.4 How to conduct performance testing? Step1:Planning

- Load (Operational) Profile
  - To better simulate the real world, not only should the load size be estimated, but also a profile of the users/activities that make up the load needs to be created.

9.4 How to conduct performance testing? Step1:Planning

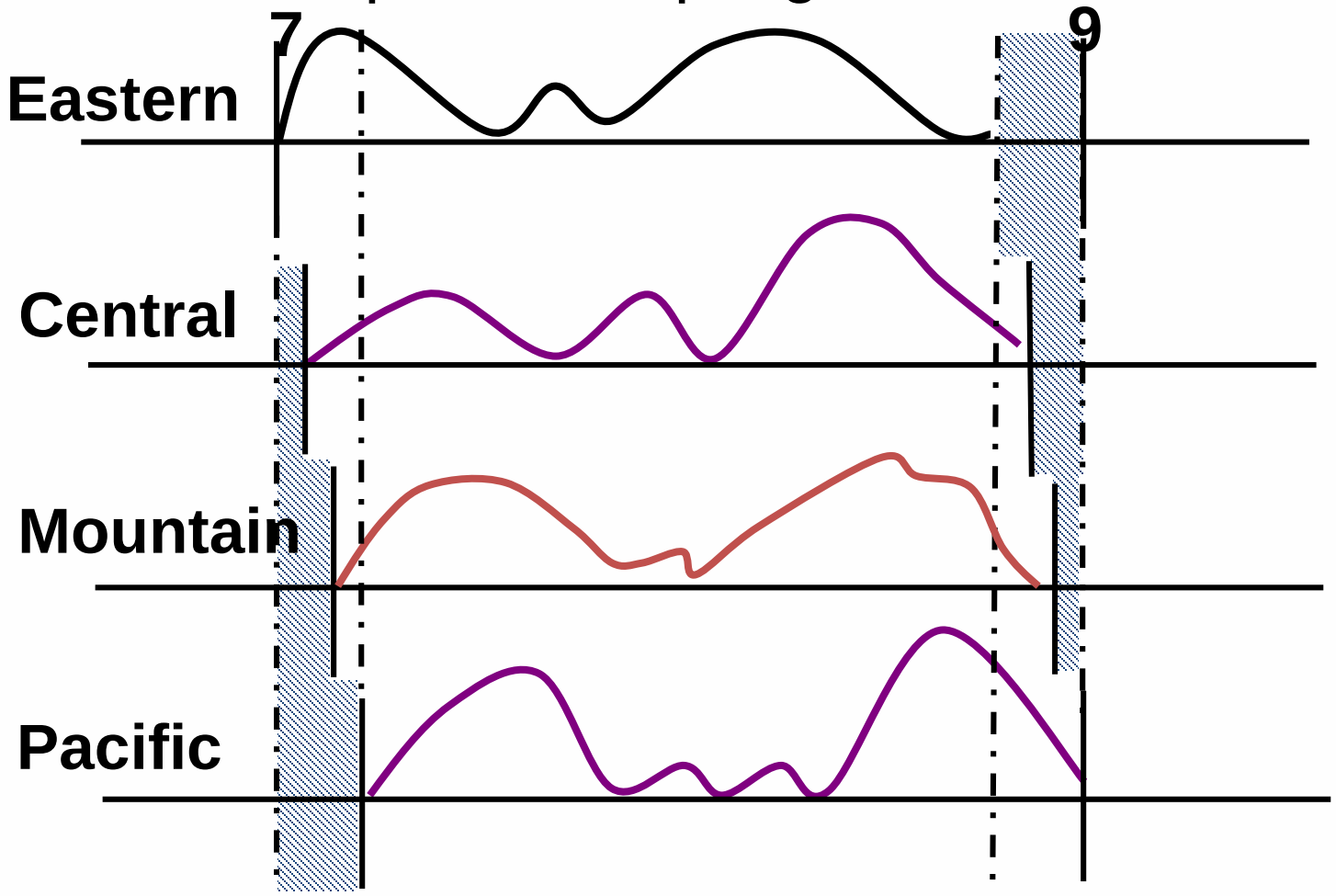
- Test Sample Sizes
  - Most factors which we want to measure, such as response time, vary based on many factors; both known and unknown
    - This means we cannot measure something once but must sample it several times in order to compute a meaningful average
    - How big does the sample need to be? This is a rough rule of thumb for the relationship between sample size and accuracy

9.4 How to conduct performance testing?
Step1:Planning

Sample Size	Accuracy (roughly)
30	3%
100	10%
200	25%
1,000	50%

9.4 How to conduct performance testing?
Step1:Planning

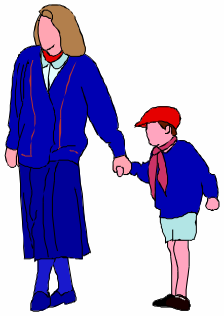
- When To Sample – Sampling Errors



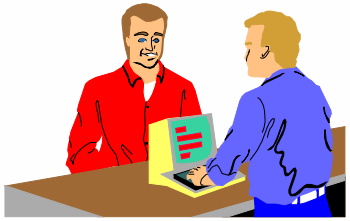
9.4 How to conduct performance testing?
Step1:Planning

- Number and types of users
  - The mix of users (events) may vary by time period.
  - Special events can alter the mix.

Browsers



Buyers

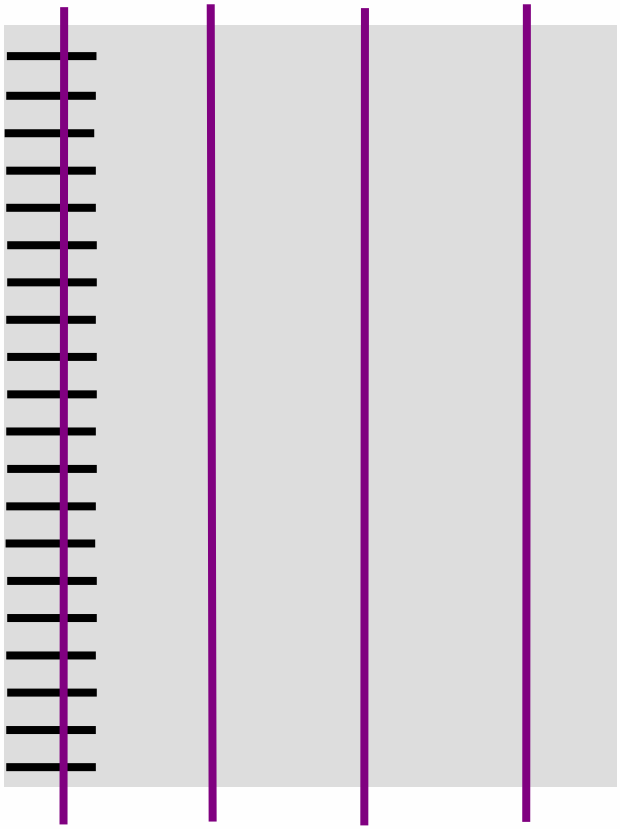


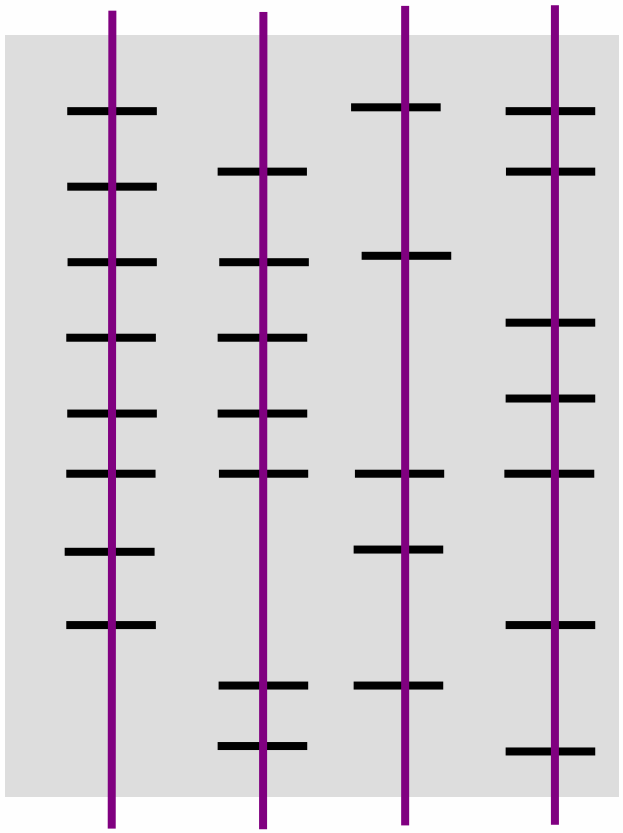
9.4 How to conduct performance testing? Step1:Planning

- The number of users on the site will vary based on many factors
  - Type of site
  - Offerings
  - Seasonal events
  - Special events
  - Etc.

9.4 How to conduct performance testing?
Step1:Planning

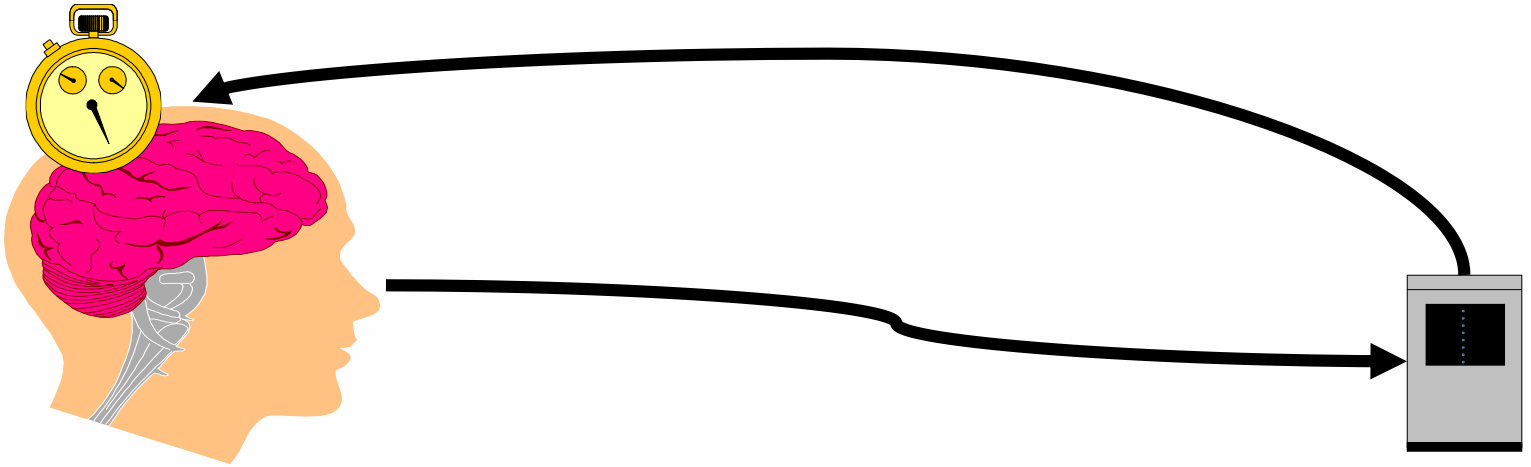
- Concurrency





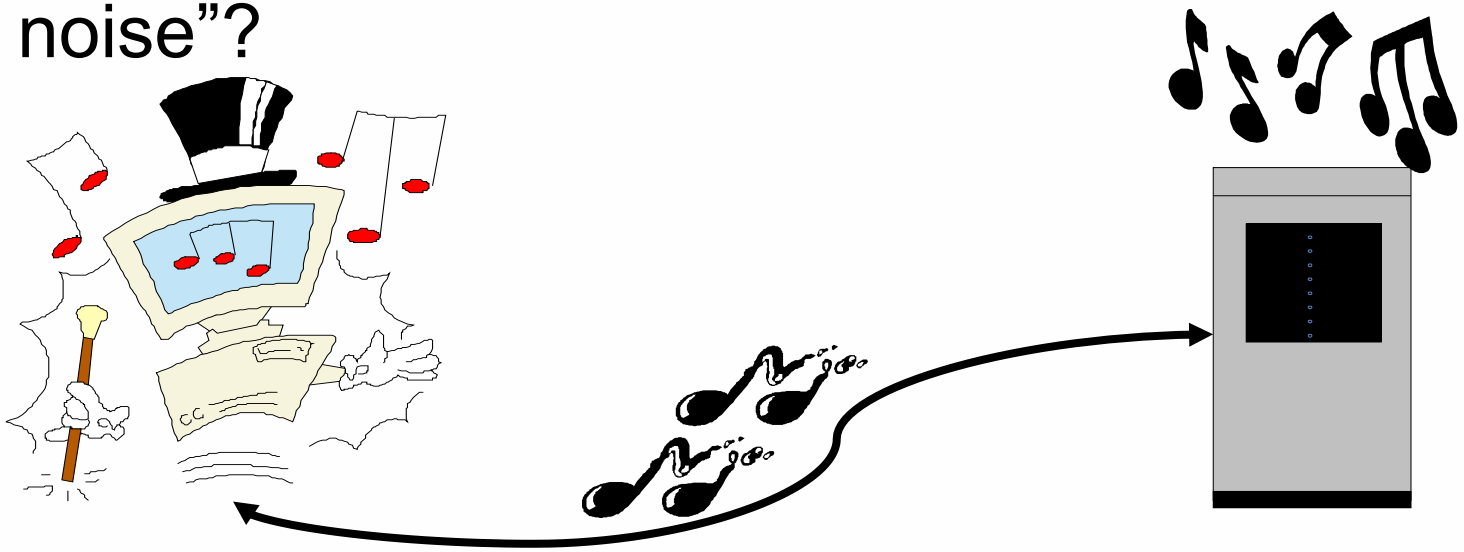
9.4 How to conduct performance testing? Step1:Planning

- Thinking time



9.4 How to conduct performance testing? Step1:Planning

- Background Load (Noise)
  - What additional activities need to be considered to accurately reflect the performance degradation caused by network and other application loads, “background noise”?



9.4 How to conduct performance testing? Step1:Planning

- Testing types and metrics needed
  - 1:Load
    - Increasing Workload
      - Start / End with how many users?
      - Increment by how many users?
      - How often to increment (Seconds)?
    - Steady-State
      - How many users?
      - How long (Days/Hours/Minutes/Seconds)?

9.4 How to conduct performance testing? Step1:Planning

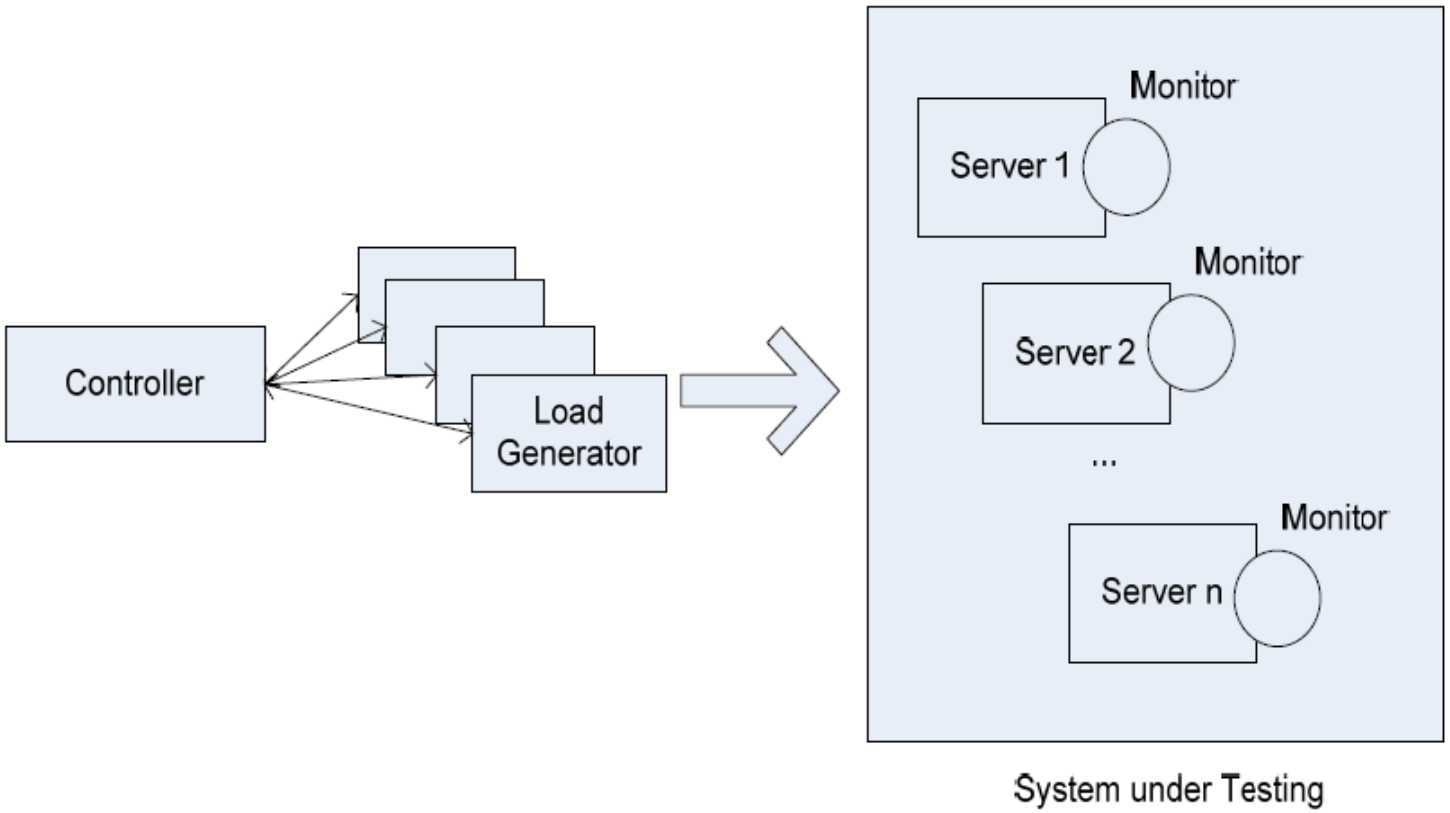
- Testing Types and Metrics needed
  - 2: Performance
    - What are the reasonable response times in seconds? (each task/page)
  - 3:Stress
    - What are the expected limits?
    - How many peak users ?

9.4 How to conduct performance testing? Step2~Step4

- The second to the fourth step：
  - Record, Modify, Execute
  - First the operation of actual users simulate behavior, record and playback multi-user test transaction processing process, automatically generating test scripts.
  - Second modify the script, increase logic control, complete the parametric and data correlation.

9.4 How to conduct performance testing? Step2~Step4

- An abstract architecture for performance testing tools



9.4 How to conduct performance testing? Step2~Step4

- A profiler（检验器） is a type of performance monitor that provides code-level measurement, including timing, memory usage, and so on.
- Profilers are useful when we want to locate performance bottlenecks in a software system.

9.4 How to conduct performance testing? Step2~Step4

- Virtual users generator
  - Used to capture the end user business processes and create test scripts, namely by recording the application in the typical end users executive of operation, record to automatically virtual users (Vuser) script, so that as the load testing foundation.
  - Virtual users include: C Vuser, VB Vuser, VB script Vuser, JAVA Vuser, Java Script Vuser……

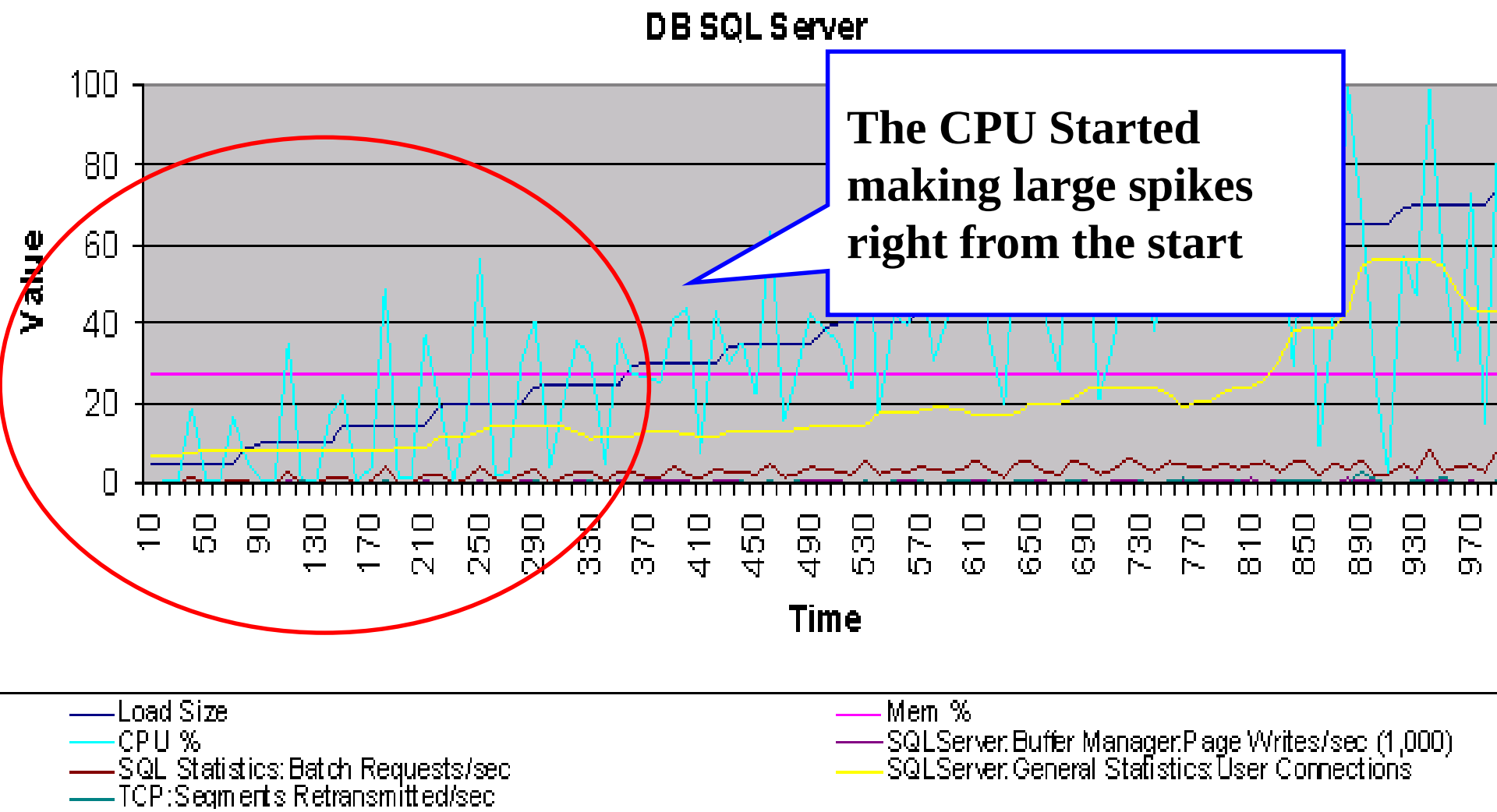
9.4 How to conduct performance testing? Step2~Step4

- Controller
  - Used to organize, driving, management and monitoring load testing central control.
  - Use controller can run used to simulate real executed operation script, and may through let multi-Vuser simultaneously execute these operations to create load in the system.

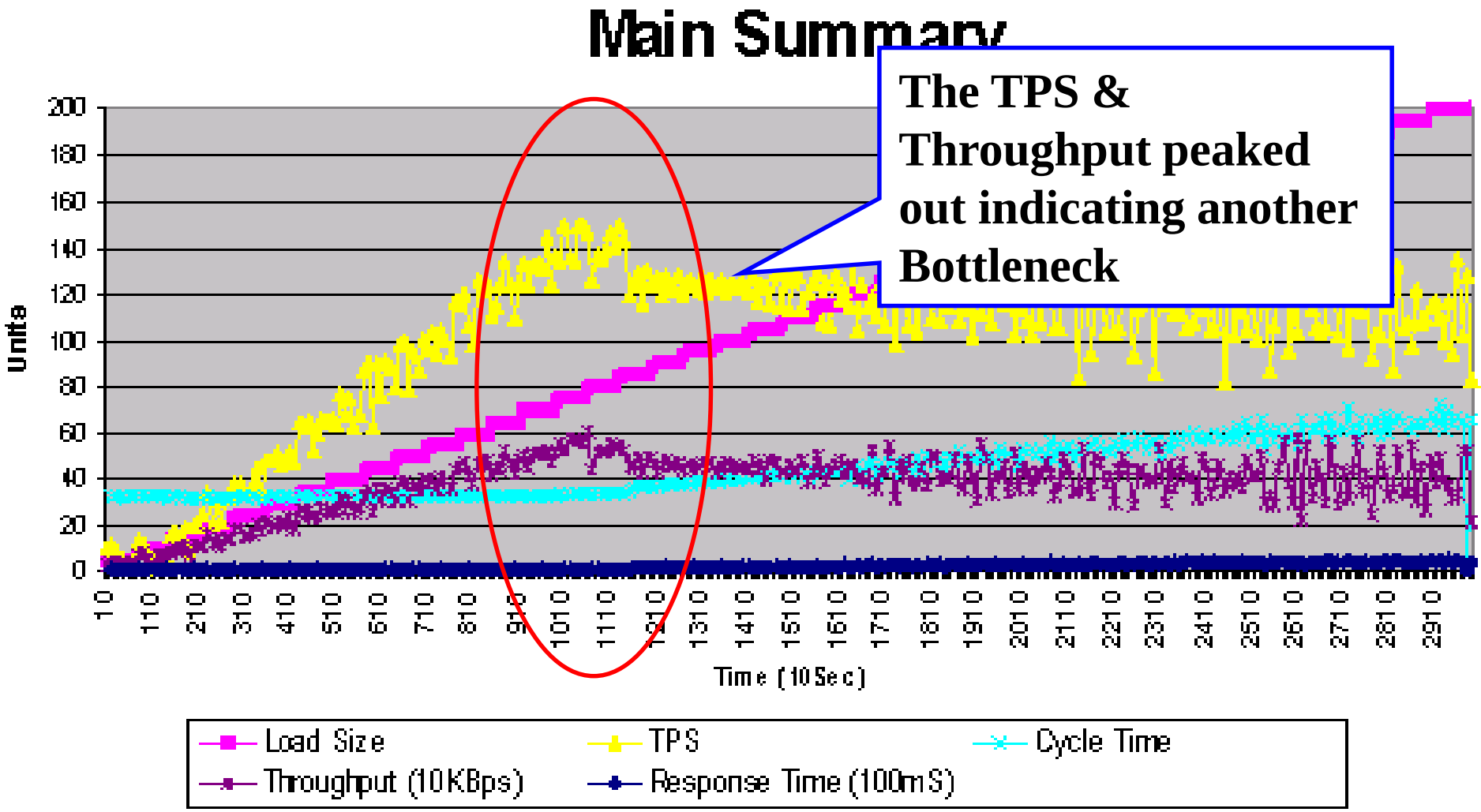
9.4 How to conduct performance testing? Step2~Step4

- Load Generator
  - Load generator to run the virtual users, to produce efficient, monitoring load.
- Analysis
  - Help to check , analysis and comparison the result of performance and use the reports and graphs to label and determine the application bottleneck, and make sure that needs to change to improve system performance of the system.

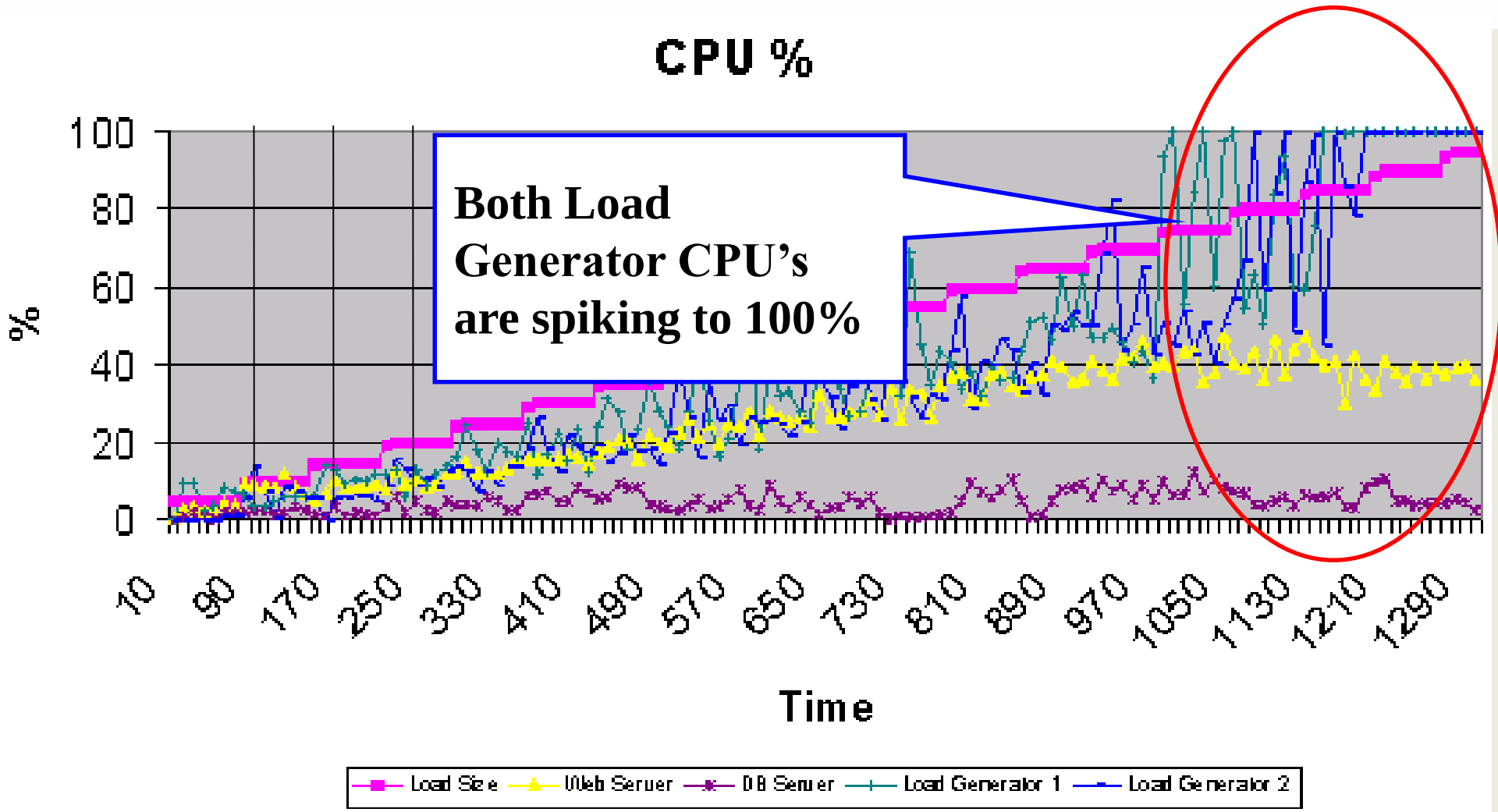
9.4 How to conduct performance testing?
Step5:Analyze



9.4 How to conduct performance testing?
Step5:Analyze



9.4 How to conduct performance testing?
Step5:Analyze



Thank you !