Jon's Performance Musings: Your 2011 Capacity Plan

Jon E. Schmidt
Transaction Design, Inc.
San Rafael, CA, 94901, USA
1.415.256.8369
inform@banbottlenecks.com

Jon is the founder of Transaction Design, Inc. (TDI), a consulting firm located in the San Francisco Area which specializes in performance and capacity studies with clients worldwide. He is the creator of the Ban Bottlenecks® service and has an extensive background in the implementation, testing, and tuning of high-availability systems.

Whew!

Another holiday season is behind us. If you support retail, hopefully your systems came through the holiday with flying colors. If you have been doing things right, you were relaxed and confident that the systems and the networks were properly sized to handle the traffic. If not, then keep reading.

One Half-Hour Per Year

That's the target for sizing your processing environment. Depending on your industry, it may be Thanksgiving (grocery), or Christmas (retail or retail banking), January flu season (pharmacy), summer drive time (gas stations), or whatever your particular industry requires. If you've sized properly for that 30-minute peak, then you can relax. If you've done your homework properly, you know what the peak demand will be, months before it hits.

Year In, Year Out

Doing that homework isn't hard. It's just work. Peaks are easy to predict, since they happen every year. Every system is affected by demand cycles. The job of a performance team is to understand those cycles and prepare for the next peak. It's not rocket science, just due diligence.

Wanted: Obsession For Perfection

So how should you create and manage a team for critical system capacity planning? My earlier comment that you need to size for one half-hour per year doesn't mean that you only have to look at the system once per year. Far from it. It takes work to properly manage a system. Business demand changes. Application software changes. Operating system software changes. The network changes. With all these changes, the team should strive for continuous improvement of the environment. Bugs need to be fixed. Operational issues need to be resolved. Every blip on the processing charts needs to be understood and corrected if necessary. Heinrich ratio issues need to be minimized. This requires a particular dedication from the team of professionals.

Capacity Team Participants

Who should be on the team?

Performance engineering: These are the folks who are capturing and examining the numbers. They need to be collecting and correlating everything, not just CPU, since modern systems rarely have a CPU problem. These folks must have some metrics from the application itself: Preferably transaction counts and response times. However, anything which is proportional to the business demand will help. They also must have network and SAN/NAS statistics. This group may be from the system engineering, capacity management, and/or operations departments.

System administration: Because they can implement system changes such as cache tuning and rebalancing.

Operations: Because they often are the "first responders" when a problem occurs. This group should also include a representative from the help desk who can talk about the customer experience and complaints.

Applications: Because they are ultimately responsible for the quality and stability of the application. They're the ones who can explain what pieces are involved in the transaction path and batch processing. This may be a vendor.

Business Line Manager: Probably the most important person in the room, because s/he is responsible to upper management for the business case for the system. This person may be the one who prepares the budgets and writes the checks for the other groups. He/she is most important because they understand the business plans.

Agenda

How should these meetings be run? I recommend a formula which works well for us and our clients:

The performance team prepares, distributes, and discusses their report: When was the prior month's peak demand, and how did the system use its resources to respond? How did that compare to the prior months' peaks? Is the system more or less efficient per transaction? What are the historical growth rates for all the major factors, especially transactions? Were there any unusual events in the month, and what caused them?

The whole team discusses the report, and identifies action items in response. They also discuss any upcoming changes to the system and environment.

The business manager discusses the plans for the business, especially such things as mergers, acquisitions, upcoming marketing promotions, new partners or new major customers, and anything else which would affect the traffic on the system over the next three to six months.

CPU and Swaps

— cpsprd — cpubz — intbz — swaps — pkswap
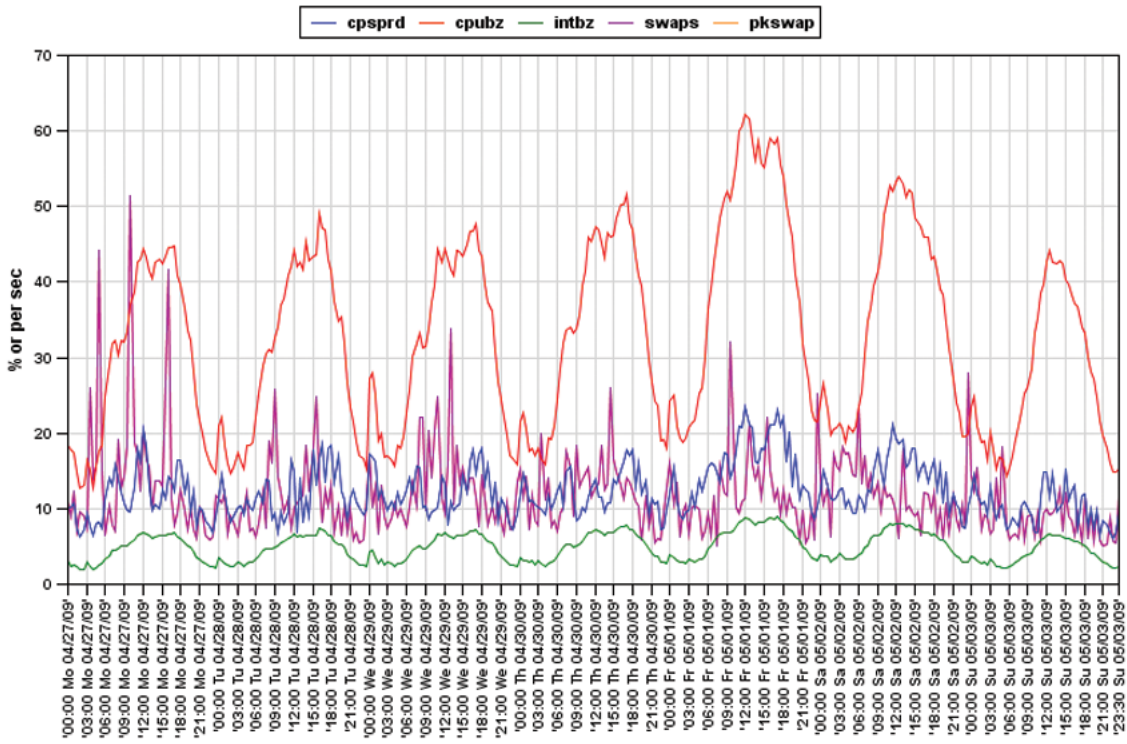
The Secret

There is one simple secret known by capacity planners: This year's ratios will probably be the same as last year's ratios. UNLESS the business has changed, which is why the business manager is the most important person in the room. The ratio of the December peak to the October peak for last year will probably be the same this year. If the usage per transaction hasn't changed, then one can take this October's numbers and apply the ratio and come up with estimates for the resources that will be needed to handle this year's traffic. And then, make sure those resources are in place comfortably before the demand hits. Then relax.