hochschule hof
University of Applied Sciences

UNIVERSIDAD DE MÁLAGA

Overview of advanced storage technologies and storage virtualization

Pablo Pérez Trabado

Dept. of Computer Architecture

University of Malaga (Spain)

Disclaimers

- Some of the figures in the slides are taken from SNIA Education tutorials. To comply with the SNIA conditions of use for these documents, any SNIA tutorial used in the elaboration of these slides has been also provided as a handout
- Many (but not all) figures in the slides have been taken from Internet-available resources, including Wikipedia. Whenever possible by copyright restrictions, the original source has been also provided as a handout along with the slides. Also, whenever possible the original source is quoted. In any case, no ownership claims are made over images in the slides not drawn by the author himself.
-
- VMware does not endorse or make any representations about third party information included in this document, nor does the inclusion of any VMware icon or diagram in this document imply such an endorsement.

hochschule hof
University of Applied Sciences

UNIVERSIDAD DE MÁLAGA

UNIVERSIDAD MÁLAGA

Introduction

Aims of course

- Improve knowledge on storage technology
  - Understand logical organization of storage I/O
  - Understand relationship with underlying physical infrastructure
  - Understand impact of storage on performance of virtual infrastructure
- Help to make better design and purchase choices, both for software and hardware

Does it matter to me?

Look, I'm going to be a database/e-commerce/multimedia IT engineer, not a SysAdm. Is storage of any importance to me?

- Today, IT firms find economic profit in handling large amounts of data:
  - High-performance databases for commercial transactions
  - Indexing/data mining over BIG databases/sets of data (*Google, oil/gas industry, etc*)
  - Multimedia contents (*YouTube*)

- Social networking is driving data growth out of control
  - Large amounts of useless stuff get also stored
  - Both valuable and useless data must be moved in/out of storage

Does it matter to me?

- Many IT engineers will need to develop applications which handle big data

- IT engineer must be aware on how his design interacts with I/O:
  - The best software system can perform poorly under badly designed I/O
  - The best/most expensive I/O design can be made run poorly by inefficient software or bad configuration

Isn’t spending more money enough?

Today we’ve got systems with bazillions cores. Won’t they handle anything the I/O can throw their way?

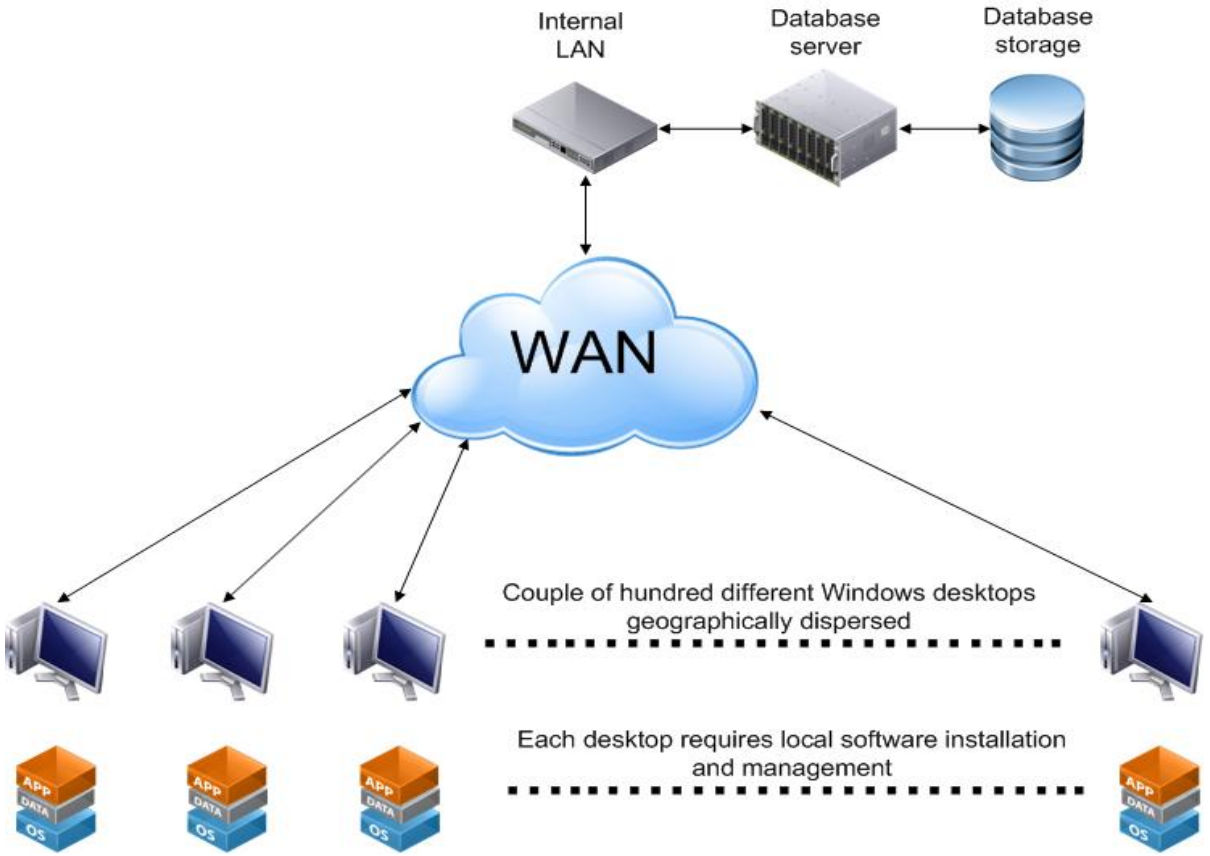
- Unfortunately, no. No matter how many CPU/GPU/cores/RAM you have, I/O bottlenecks can (and will) bring even the most powerful supercomputer to its knees:
  - Processing will stall, and CPUs idle
  - Costs (lots of) money
    - Waste of leased CPU time
    - Waste of electric power and cooling costs
    - Waste of investment in processing power (poor ROI)

Isn’t spending more money enough?

Disks are dirty cheap nowadays. If you’ve got I/O problems, just buy larger/faster storage and you’ll be OK, right?

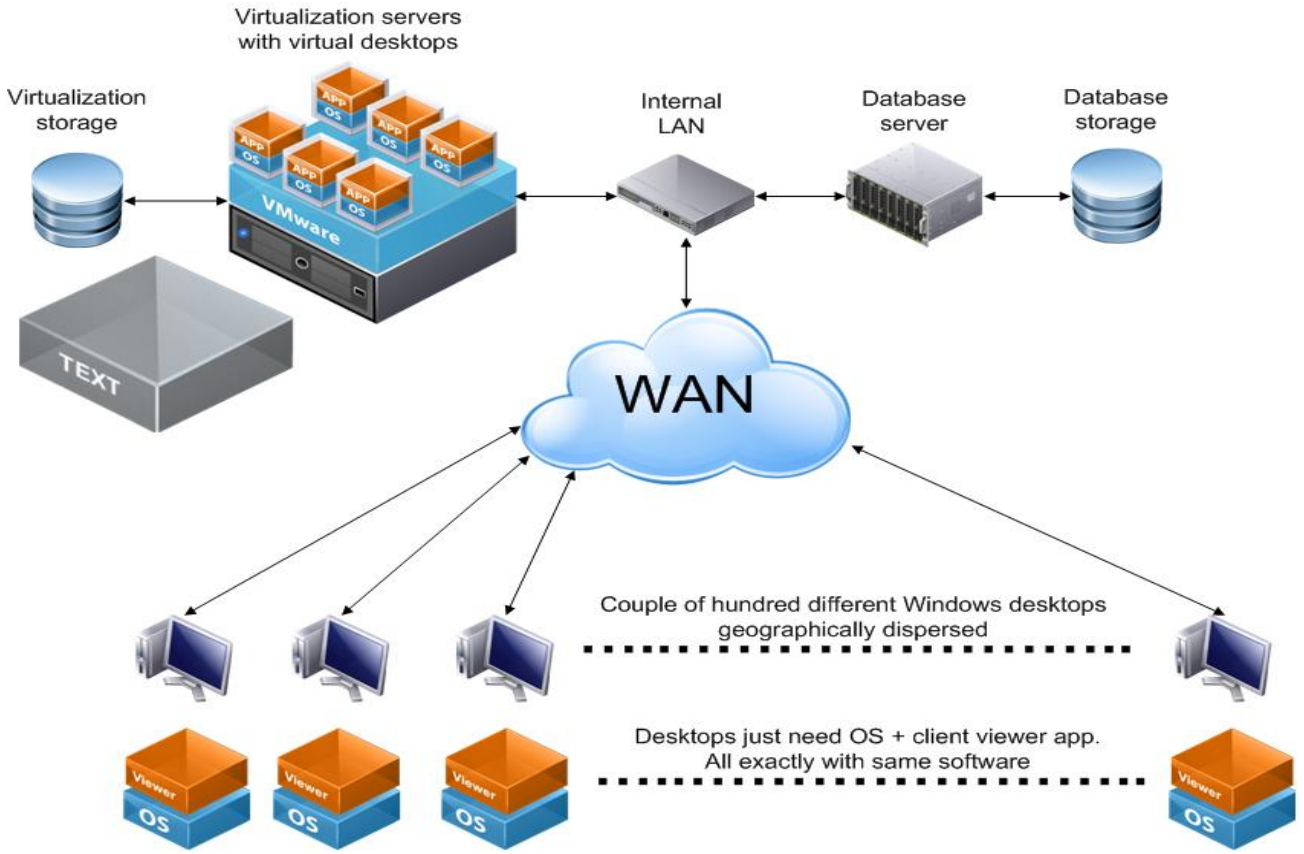
- Not necessarily. Without a clear understanding of the reasons of the bottleneck, you may end just wasting…
  - the firm’s money (lots of it!)
  - the staff time (productivity loss)
  - everyone’s patience with you !!!

Example



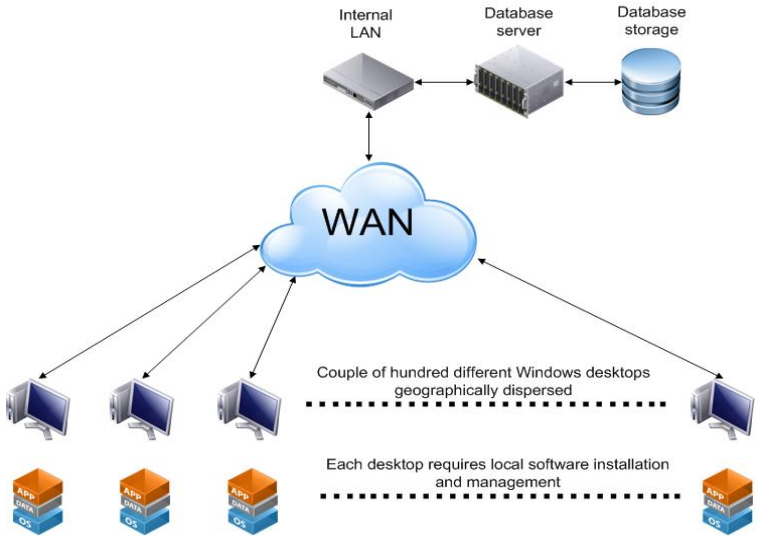
- Firm's management app must be accessed from many separate physical locations
- Each location requires desktop, with OS + App software installed
- High costs of software maintenance

Example



- Under your advice, firm introduces VDI (*Virtual Desktop Infrastructure*)
- Desktops with App run as virtual machines in centralized server
- Physical desktops now are just remote viewing clients
  - Great savings in Opex (Operational Expenses)

Example



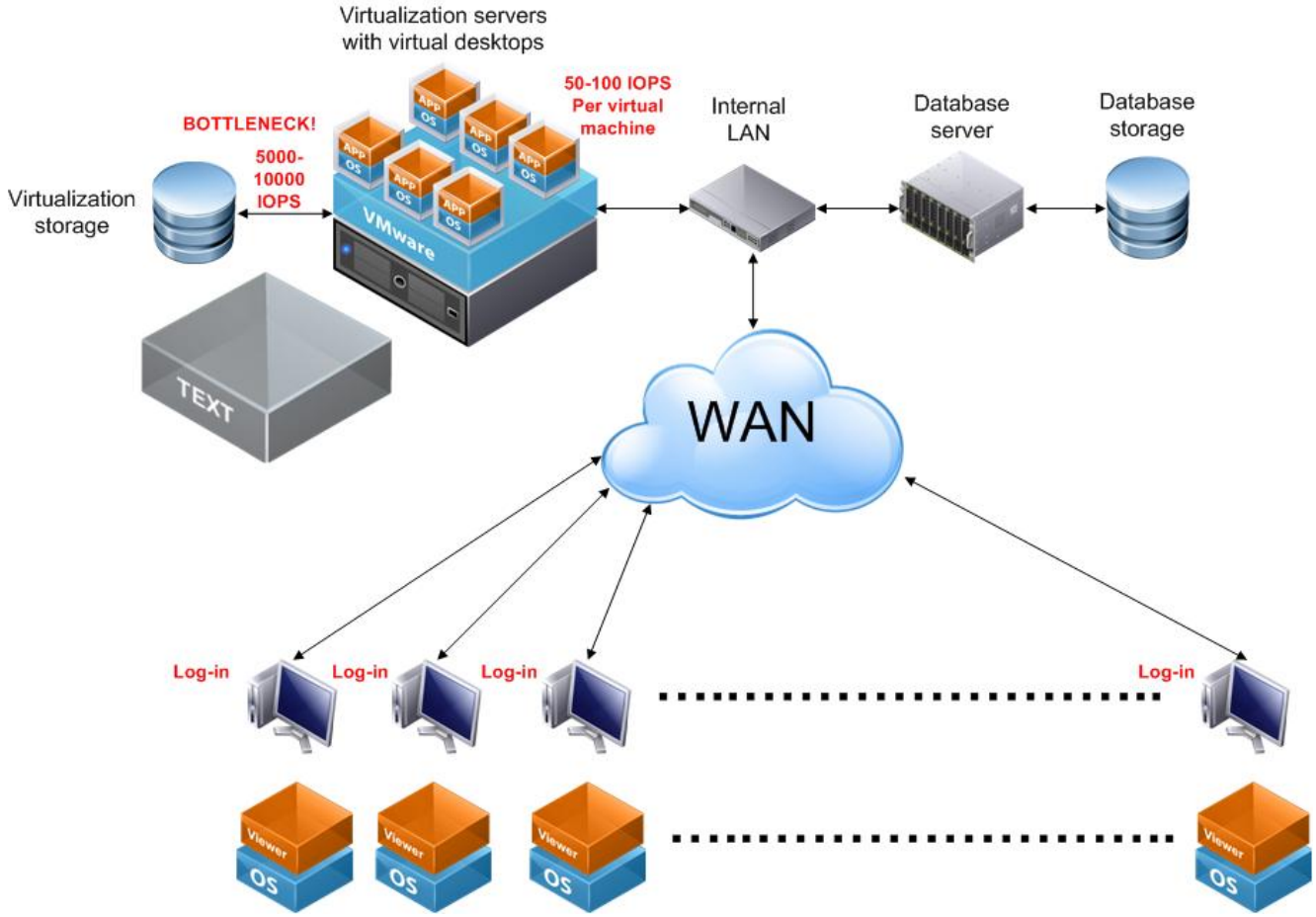
- System turns out to work very well… <u>almost</u>
- Performance of system sinks around opening hour (9:00 am)
  - Login process slows down to a crawl
  - Login and app startup can take up to 5-10 minutes to complete
  - Users can not do any work until login is complete
  - Users get frustrated and angry with new system
  - Users complain to boss
  - Boss gets angry with ***<u>you</u>***

Example

- Your problem is called "*VDI boot storm*"

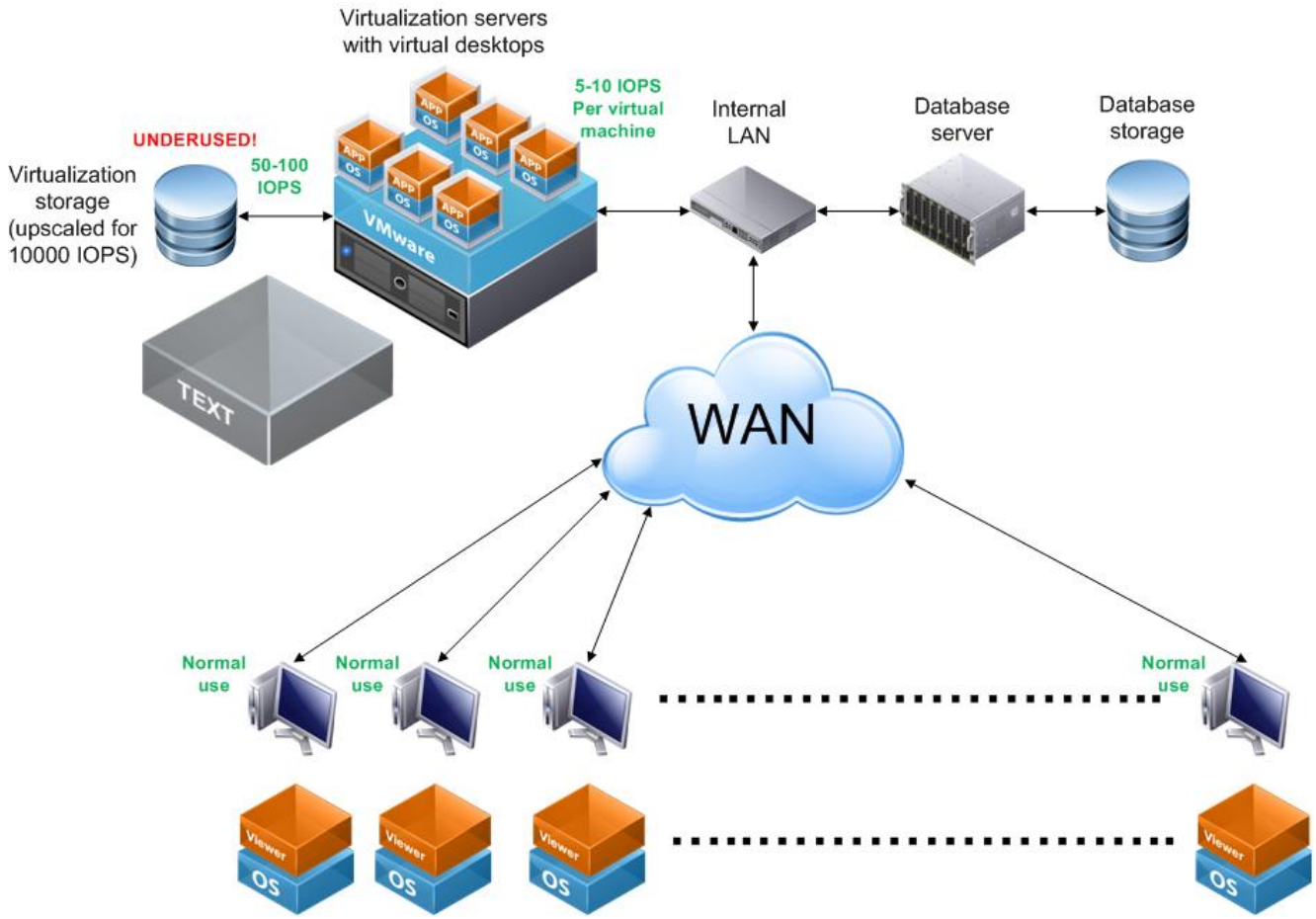
- Workstation login process requires lots of disk access
  - IOPS = I/O Operations Per Second
  - In normal use, workstation generates just 5-10 IOPS
  - While booting, workstation issues 50-100 IOPS

Example



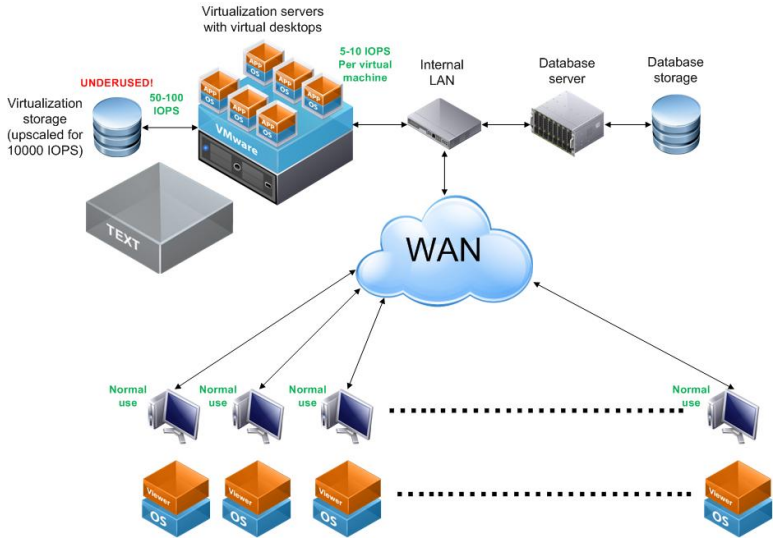
- Virtualization has created a bottleneck, centralizing all IOPS into a single storage appliance

Example



- You solve the problem by brute force, purchasing new storage
  - More disks ("spindles"), faster (15k), much more expensive
  - Probably need to change also disk rack

Example



- Problem: after login is finished, App uses storage at just 1% of possible IOPS
  - System overdimensioned for normal use
  - But power and maintenance costs must be paid for all that excess capacity
  - Storage costs negate any savings from reduced software maintenance
  - Bad ROI (Return Of Inversion)
  - Boss unhappy with ***you***