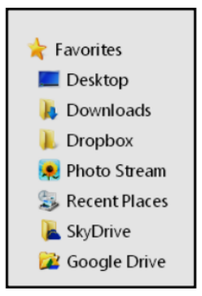
Week 6: Cloud storage

- It is now common for people to have their data saved 'to the cloud'.
- It is typically implemented as
  – A website where upload / download is a browser activity
  – A 'virtual peripheral' or device, or
    • In other words, it can be made to look like any other media storage option.
- Potentially massive storage is possible
  – All transaction data as they happen. Anywhere. eg. Security video feeds for 20 cameras around your property for a month



Cloud storage business model

Business Model – Why are they so nice to do this for you?

– Customer lock-in
  • They become a part of your business, hopefully one that become essential to you and so guarantees ongoing business

– They get to look at your data
  • Do your customers know this? Should they?
  • Google's search engine optimisation only works if the data volume feeding it is massive

– It opens other possibilities for business for them. For you also.
  • Offline Backup
  • Data Analytics, Particularly allowing you to 'compare' your data with other people's data even if you don't have access to the other people's data itself

– It facilitates project collaboration
  • within your business
    – geographic diversity is possible
    – your business can be lean and mean – and thin.
  • with outside service providers
    – Stick to what you know well,
      » and pay others for what they know well.

– It makes out-sourcing easier, cheaper since forms of access are uniform
  • With cloud providers as facilitators, job/people providers
  • Common data and processing formats
    – Accounts, Payroll, Spare parts, Delivery Tracking,

Cloud storage – risks

- Advantages
  – Simplifying your system
    • Backup is no longer your problem
    • The data is available on all your devices
      – interfacing is no longer your problem
  – Sharing
  • The data can be easily shared with others at user level
  • Some vendors give a precise level of control
  • Security break-in issues are 'not your problem'
- Disadvantages
  – Speed, limited by network bandwidth
    • How fast can you access your transaction data
  – Security
    • What do they know about your business and customers?
      – What do you know about them ?
  • You are legally responsible for the security of your data – not them.
  • Impact of failure
    – If your data gets lost, what happens to them? What happens to you!!
    – How much is your data worth?
      » Can you prove it?
      » Can you recover / regenerate it?

As always: *Look at the terms of service for answers*

Big data

In some ways, a special case of cloud data storage,

- Characterised by the "3,4,5 or 7 V's"
  – **Volume**
    • It's BIG. TB / PB per day are not unknown
  – eg. internet of things (IoT) device data streams
  – **Velocity**
    • generated very quickly – typically it is streaming data
      – no time for detailed analysis directly on the data itself
      – best to quickly summarize and analyse these instead
  – **Volatility / Variability**
    • original streaming data not stored – only summary or analytics
    • analysis cannot be repeated if wrong
  – **Veracity / Validity**
    • you have to trust it, since there is no time to double-check
  – **Variety**
    • data formats are often not uniform, may need meta-analysis
  – **[Value, Visualisation]**
- ***Who is responsible for the consequences of bad data?***