Amazon Web Services

Monday, July 4, 2016

11:12 PM

**Cloud Computing / AWS:**

* Provide compute, storage and database services for businesses to run applications
  + Pay-as-you-go --> little capital investment
  + Worldwide network of data centres --> customer enjoys fast services anywhere
* Why customers choose AWS vs. build their own data centres?
  + Flexible operations: easily upsize/downsize/move/expand business to respond to changes
  + Turn fixed costs to variable costs --> lower breakeven point --> lower operational risk
  + Benefit from economy of scale --> lower operational costs
  + Lower risk to try out new initiatives --> encourage innovation
* Why AWS is a sustainable business model?
  + Trend: global market, constantly changing environment --> flexibility is key for large corporations to maintain their competitive advantage
  + Small businesses don't have the resources and capacity to build and maintain their own data centres

**Internal operations - Infrastructure:**

* Amazon has x number of servers: boxes (more powerful server = bigger box)
* Amazon customers purchase y number of packages: cubes (more expensive package = bigger cube)
* Amazon's promise to customers: The boxes could hold all cubes at all times (99.95% uptime every month)
* Amazon's internal goal: Maximum capacity - the boxes are full at all times

**How we do this?**

* Capacity planning: estimate customer demand --> provision appropriate number of machines to meet customer needsWhat
  + Number of servers
  + Capacity of servers (i.e. stronger servers)
  + Distribution of servers in different geographical regions / data centres
* Maintain operations as expected/planned
  + Failures of servers - repair/replace
  + Plan hardware spend for infrastructure fleet (servers)
    - Ex: Maintain 36 Celsius degree = 1 server fails/day VS. Maintain 37 Celsius degree = 10 servers fail/day --> costs of AC/fan vs. costs of repair/replace (consider qualitative factors such as customer experience during downtime)
    - Scenario analysis (short-term and long-term costs, quantitative and qualitative factors) to make decision
    - Breakdown of costs by vendors (HP vs. Dell) --> vendor evaluation
    - Track revenues and expenses for each data centre/geographical region
    - Peak time vs. slow time (customer usage)
    - Behaviour of different types of operating costs (i.e. electricity, internet) - relationship between costs and level of activity
      * Variable costs vs. fixed costs vs. step fixed costs
      * Special characteristics: the 90h percentile pricing model for internet
* Minimize idle resources:
  + Promotions for slow periods --> more sales (positive incremental revenue)
  + Perks (i.e. serve more than what's included in the package) to existing customers during slow time --> increase customer experience
    - Let them know so that they don't expect this on a regular basis
    - Focus on customers with high potentials
  + Efficiency of operations (% usage vs. % idle through time)
* Project management for development of new tools:
  + Financial planning for proposal of new projects (forecasted costs, timeline, expected benefits)
  + Track progress (timeline) and costs
  + Measure benefits during and after development

**Support Corporate IT Engineering headcount?**