Risks

Wednesday, March 17, 2010
11:24 AM

Risks

Main business case for clouds: assuming risks

Key risks:

Availability: will I be able to use the site?

Response: how fast?

Privacy: will data be readable by unauthorized parties?

Security: will data be editable by unauthorized parties?

Contingency: If the worst happens, will data survive?

Availability

Wednesday, March 17, 2010
11:37 AM

Availability

Risk: site will not be available.

Cost: losses due to downtime (Patterson's formula).

Mitigation: redundancy.

Duplicated data: lose one server, doesn't lose data.

Duplicated infrastructure: geographical

Response

Wednesday, March 17, 2010
11:38 AM

Response time:

Risk: delays in response time.

Cost: loss of work/business.

Mitigation: scalability and elasticity.

Wednesday, March 17, 2010
1:11 PM

Privacy

Risk: exposure of private data of business or individuals.

Cost: lawsuits, reputation, loss of competitive advantage, Federal prosecution for non-compliance.

Mitigation: security best practices, data retention policies.

Security

Wednesday, March 17, 2010
1:13 PM

Security:

Risk: critical data is corrupted.

Cost: reputation/customers, legal action, Federal prosecution for non-compliance.

Mitigation: security best practices.

Contingency

Risk: hardware rot, natural disaster, etc.

Cost: downtime, loss of data integrity.

Mitigation: data backups, redundant storage.

Myth and reality

A common myth: cost of providing service is the cost of setup and power.

The reality: most of the cost of providing service is risk mitigation.

Human staff is main cost and bounded resource.

The slippery concept of risk

Beattie et al, "Timing the application of security patches for optimal uptime", Proc. LISA 2002.
http://www.usenix.org/event/lisa02/tech/beattie.html

If uptime is important, then one must balance the risk of **work lost due to patches** against the risk of **security violations**.

Patching early leads to downtime and work lost; one must often "patch the patch."

In some cases, optimal strategy is to **wait a month** before patching a security hole, to **balance risk of downtime against risk of infection.**

Basics of risk analysis

Wednesday, March 17, 2010
1:47 PM

Risk exposure

= expected value of a risk

= ∑(Prob(outcome)*(cost of outcome))

where Prob(outcome) is the estimated probability of the outcome, for **mutually exclusive outcomes.**

(cost of outcome) is how much the outcome costs.

Game: **minimize exposure** by **changing probabilities or costs.**

But not the whole picture:

Wednesday, March 17, 2010
3:00 PM

But not the whole picture:

Risk has a time component.

A **trial** is one-time interval in which a risk event may occur or not, e.g., one week.

However, cost of mitigation have a different time component, e.g., a yearly contract.

The previous equation covers **one trial.**

(end of lecture on 4/13/2011)

An alternative formulation

Wednesday, March 17, 2010
2:56 PM

An alternative formulation

Incidents arrive at a Poisson rate λ.

Each incident has an average cost C.

→ Exposure in interval ΔT = λ*ΔT*C.

Why it is equivalent: the Poisson postulates

Prob(1 event in ΔT) = λΔT (for ΔT small enough)

A curious example: the business case for anti-viruses

Wednesday, March 17, 2010
1:56 PM

The business case for anti-viruses

What business value does an anti-virus have?

Risk: viral infection.

Impact/cost: denial of service or illicit use of business infrastructure (e.g., for bots).

Trial: time between virus outbreak and definition update. No such thing as no exposure.

Can compute average arrival rate for viruses from history (see CERT statistics).

Can compute average impact (e.g., desktops affected) per incident.

Exposure without anti-virus

Propogation/incident determined by when incident is discovered, and percolation.

Average downtime/incident D based upon human response time = wait for help + time to solution, which is a function of the # of humans n available to help and propogation (i.e., how many stations affected).

Average cost/downtime C.

High arrival rate λ for incidents.

Exposure in $\Delta T = \lambda \Delta T C D$.

Computing D

Suppose there are L humans to walk around to fix viruses.

Little's laws: requests in system L = mean arrival rate λ * mean time in system W. W=L/λ.

Obviously, "how bad things get" depends upon arrival rate λ of incidents.

(in reality, incursions can quickly progress beyond steady-state).

Exposure with anti-virus

- No such thing as 0 exposure.
- Race between virus creation and signature posting.
- Risk is that virus will arrive between creation time and rule posting time.
- Decompose λ into $\lambda_1 + \lambda_2$, where λ_1 is arrivals between creation and posting, and λ_2 is arrivals after posting.
- $\lambda_1 << \lambda_2$.
- Exposure = $\lambda_1 \Delta T\ C\ D << \lambda \Delta T\ C\ D$.
- Value of anti-virus = $\lambda_2 \Delta T\ C\ D$.

Risk-aversion

Some costs are infinite; i.e., **business goes bankrupt.**

Example: massive privacy violation; trust lost.

Mitigation strategy: best available.

Acceptable risks

Some risks are acceptable, i.e., downtime due to disk failure.

But risks change over time.

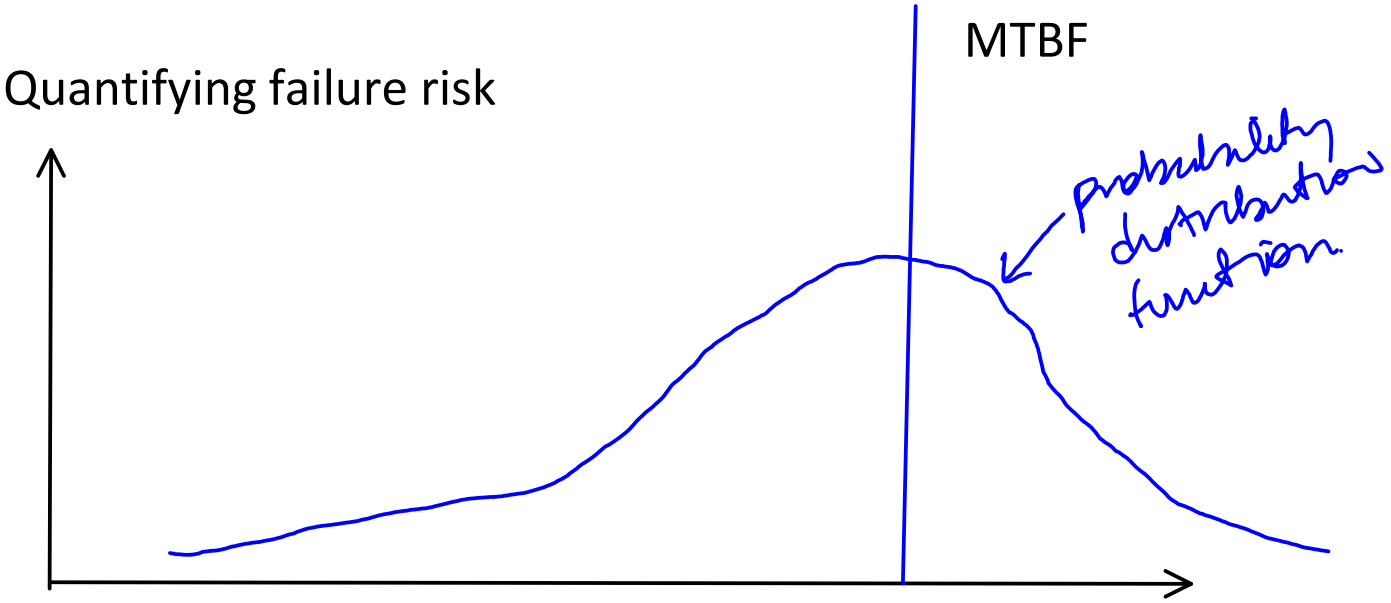
Example: MTBF (Mean Time Before Failure).

Distribution of disk failures around MTBF time.

As time passes, failure of an individual disk becomes more likely.

Quantifying failure risk

Wednesday, March 17, 2010
4:52 PM



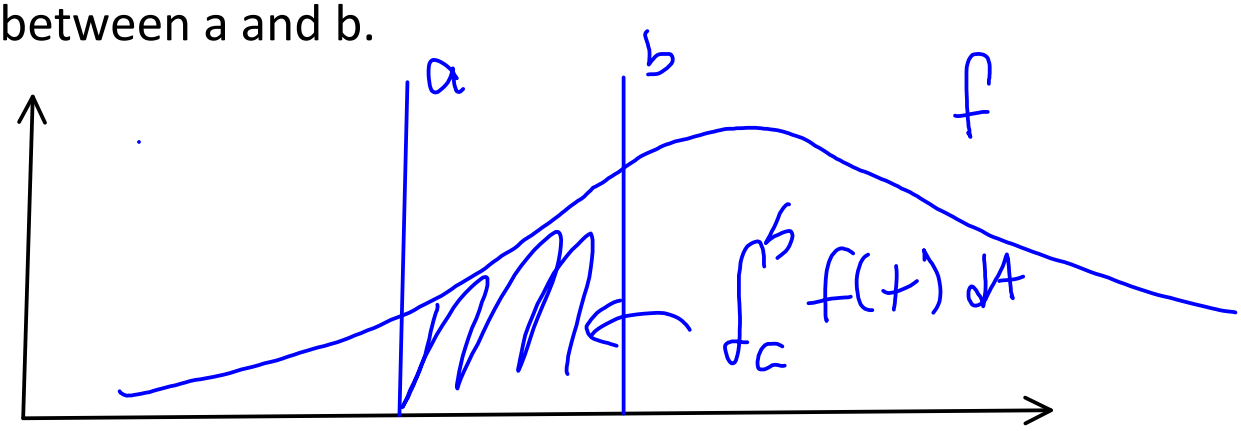
Recall: probability distribution functions

Wednesday, March 17, 2010
4:56 PM

Probability distribution function f(t):

integral of f over all t is 1.0

integral of f from a to b is the probability of a failure between a and b.



Disk risk

Wednesday, March 17, 2010
4:59 PM

Disk failure risk

Changes with time spent running.

Prob(failure within r hours) =

$$\int_0^r f(t)\,dt$$

But what about RAID?

Wednesday, March 17, 2010
5:01 PM

But what about RAID?

- Point of RAID: one failure in 5 doesn't cause data loss/downtime
- So, we're more concerned with whether 2 fail at the same time (independent events) than whether one fails.
- Downtime for 1 failure=0 (response time changes)
- Downtime for 2 failures= substantive (recover from backup or image).

A very counter-intuitive fact

- Disks age on average at roughly the same rate so that
- It is a **bad idea to deploy a batch at the same time** because
- They'll all fail at roughly the same time.
- Human labor is required to replace a disk.
- Human labor is a **bounded resource.**
- Reason: the human part of the risk equation.
- Optimal strategy: trickle-deployment, renewing disks one at a time.

The case for cloud storage

Wednesday, March 17, 2010
5:19 PM

The risk-based case for cloud storage:

"Pay someone else" to

- redundantly store data.
- scale up in response to load.
- replace and phase in disks.
- recover from disk failures.
- maintain an inventory of replacements.

versus:

- Retain human staff to replace disks.
- Keep your own inventory.
- Make your own backups.
- Create your own data storage policy.
- Etc.

Why clouds exist:

- The main providers were doing this already!