



Apdex Symposium, Las Vegas  
CMG Session 358

Discussion: is  
Apdex sharp  
enough?

Alan Ackers, Macro 4

▶ Discussion: is Apdex sharp enough?



### Introduction and background (Hidden slide)

- Macro 4
- My role
- UKCMG
- Objective of session to stimulate discussion

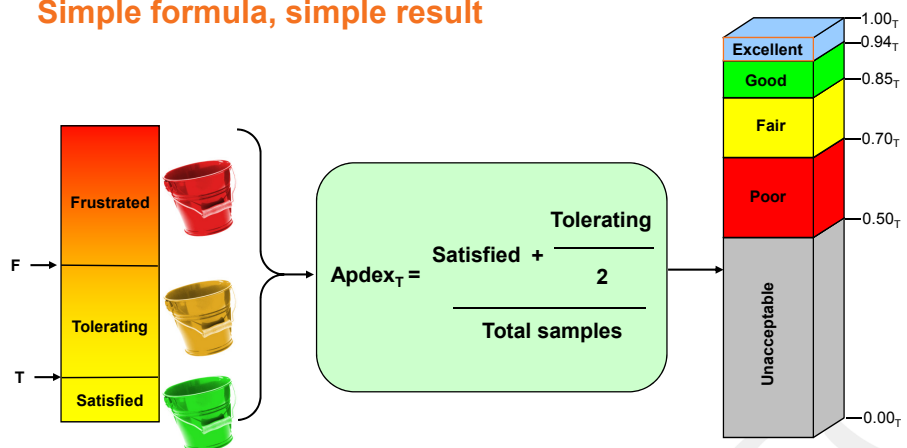


### Apdex current state

- Intended use in end user experience reporting
- 100+ measurement samples
- 3 measurement “buckets”
  - Satisfied
  - Tolerating
  - Frustrated



### Simple formula, simple result



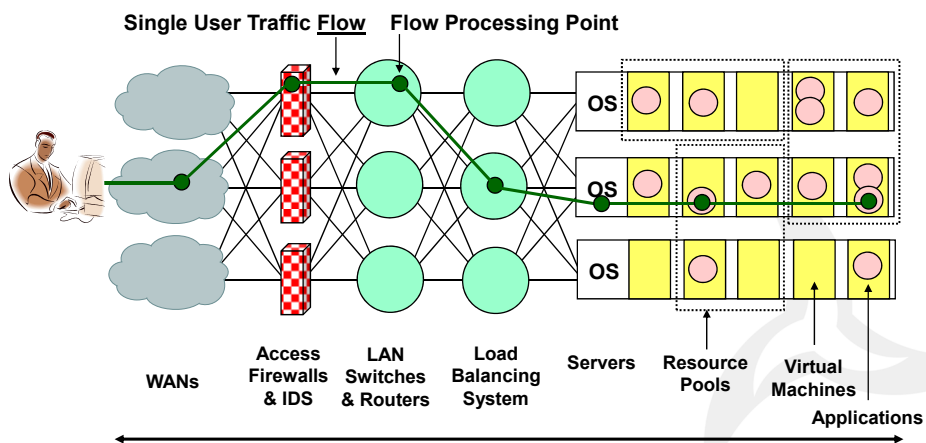
0 to 1 - simple

### But what if...?

- All the samples are grouped into a very tight time range
- Example (worst case scenario!)
  - 100 samples, T=4s, F=16s
  - Case 1: all samples 4.01s  
$$\text{Apdex} = (0 + (100/2)) / 100 = 0.5_4$$
  - Case 2: all samples 15.99s  
$$\text{Apdex} = (0 + (100/2)) / 100 = 0.5_4$$
  - Both give same result...

### Is this extreme case likely?

- Not at all, in end user experience scenarios (too much potential for variability)



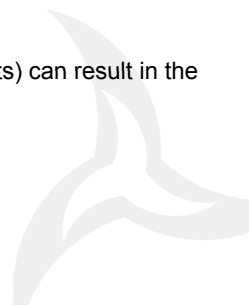
### Choice of T is key

- Low value
  - entire long tail curve is in the tolerating or frustrated bucket
- High value
  - almost all the samples fall within the satisfied bucket



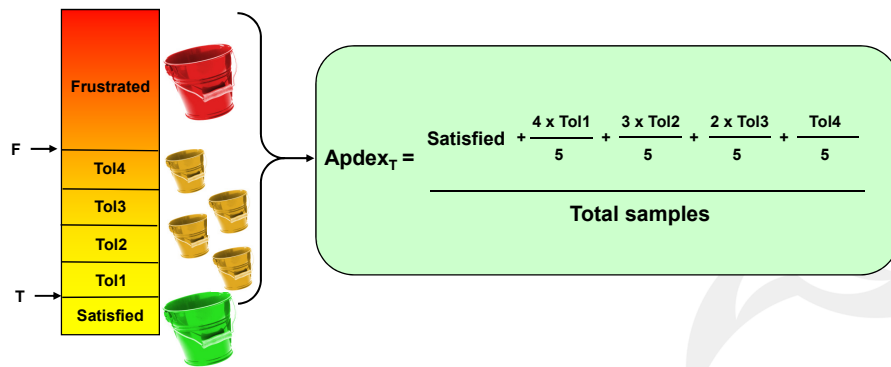
### However...

- Increasingly, the simplicity of Apdex is being applied to other situations, which can make grouping more likely
- For example, Apdex being applied to a well-defined and limited transaction such as a database lookup
- In this scenario, measurement accuracy (rounding effects) can result in the same response time for many transactions



### Possible “sharpening” – option 1

- More buckets
- For example, divide the tolerating band into 4 buckets:



### How does this change our theoretical worst case scenario?

- All the samples are grouped into a very tight time range
  - 100 samples, T=4s, F=16s
  - Case 1: all samples 4.01s (Tol1 bucket)
 
$$\text{Apdex} = (0 + ((4 \times 100)/5) + 0 + 0 + 0) / 100 = 0.8_4 \quad \text{F}$$
  - Case 2: all samples 15.99s (Tol4 bucket)
 
$$\text{Apdex} = (0 + 0 + 0 + 0 + (100/5)) / 100 = 0.2_4 \quad \text{U}$$
  - Much better

### Possible “sharpening” – option 1 (more buckets)

- Pros
  - Sharper
- Cons
  - Not as simple
  - The sample sorting process is far more involved (twice as many buckets and comparisons needed)
  - The calculation is far more complicated
  - Not good for tools that need to calculate Apdex in real time
  - Additional CPU usage for tools
  - Slower calculations

### Possible “sharpening” – option 2

- Allow the relationship between F and T to vary
- Therefore, F can be set independently
- This allows the tolerating band to be narrowed or widened depending on the situation
  
- In our theoretical worst case scenario, this would allow a lower value of F to be set (for example, 8s). Therefore, in the case of the 15.99s samples, all of these would fall into the frustrated band and receive no Apdex “credit”.

### Possible “sharpening” – option 2 (can vary F)

- Pros
  - More flexibility
  - The sorting and calculation processes are still just as simple
- Cons
  - Not quite as simple as before as F must be decided too
  - Presentation of Apdex values must specify the F values as well as T
  - The 4 x relationship is a useful relationship (this would be lost)



Discussion: is  
Apdex sharp  
enough?

Alan Ackers, Macro 4