"The Ball" Radical Cloud Resource Consolidation

Kalle Launiala, ProtonIT, "The Ball"

kalle.launiala@protonit.net, +358445575665

What to Consolidate?

(Cloud) Infrastructure Resource Usage => Reduce Costs

Primitives That Cost

 Pricing calculator ("vendor agnostic"): <u>http://www.windowsazure.com/en-us/pricing/calculator/</u>

Stand-by Costs

Storage = \$0.07 per GB per month

Action Costs

- Storage Transactions = \$0.01 per 100000
- Outbound Network = \$0.12 per GB
- Computing = \$0.02 \$1.80 per hour

Bound by realities

- Storage = maintained, on-line, scalability, capacity
- Network = maintained capacity, bandwidth
- Computing = active virtualized hardware
 - Flavors in memory, amount of cores
- Compared to existing resources, it feels expensive
 - Include maintenance troubles/costs, and it's not

What can I get for \$100 USD?

- 1429 GB of storage
 - For one month
- 1,000,000,000 storage transactions
- 833 GB of network outbound storage
- 55 to 5000 hours of computing
 - 2.3 days to 6.6 months

Computing is the target

The most expensive component...

What is computing?

What is computing?

Computer/VM that is powered up and running

The costs is to pay for reserved capacity

What if I don't actively compute anything?

Online Services = Computing

Servers serving web pages

Database servers responding to queries

What about the time when they're idle?

- Cost to pay comes from available/reserved capacity
 - Tools are estimating and reactive auto-scaling

How much computing is needed?

- Cheapest/smallest instance is \$15 USD a month
 - 768 MB of RAM, shared CPU core
- That's not much for anything real-time responses
- It still has quite a bit of CPU cycles during the month, much more than average office server uses for its tasks combined for that month.
- What if... Computing could be counted as CPU cycles?

Reminding the Ball model

- Pure distributed model to security groups'
 - Platform/instance/installation level owns no data
 - Infrastructure services come from the core

Each group owns the information

- Processing is ran within the group's context
 - ... Infrastructure such as network / storage is provided

Let's see some numbers

"The Ball" unified processing with real website group

The Scenario: Company Website

• 30 content pages

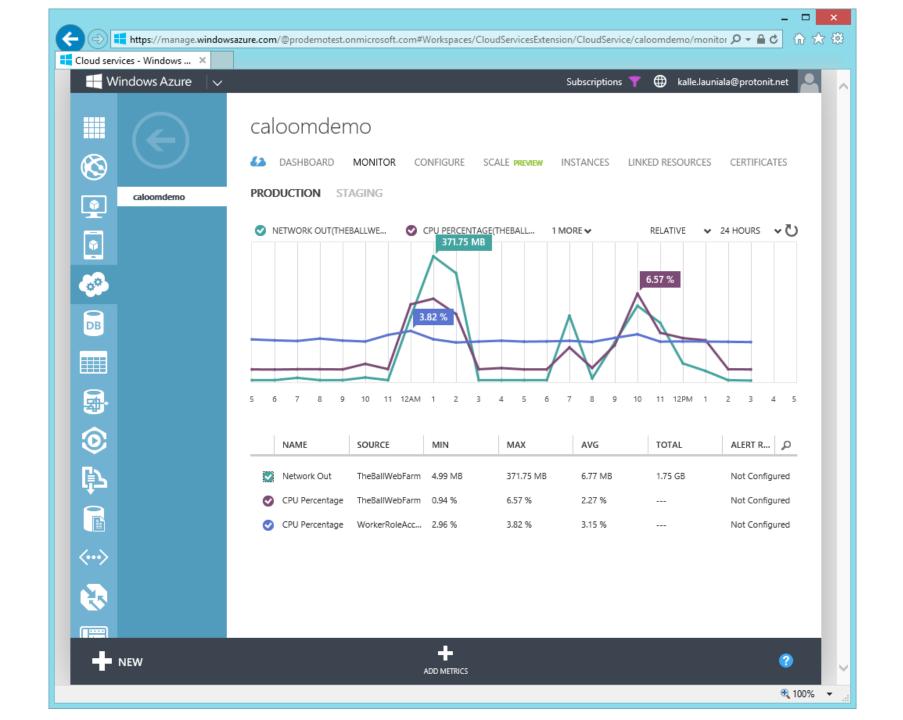
• 300 edits / month to maintain them

- 1500 visitors a month (50 a day)
 - Each land to home page and click 8-9 articles

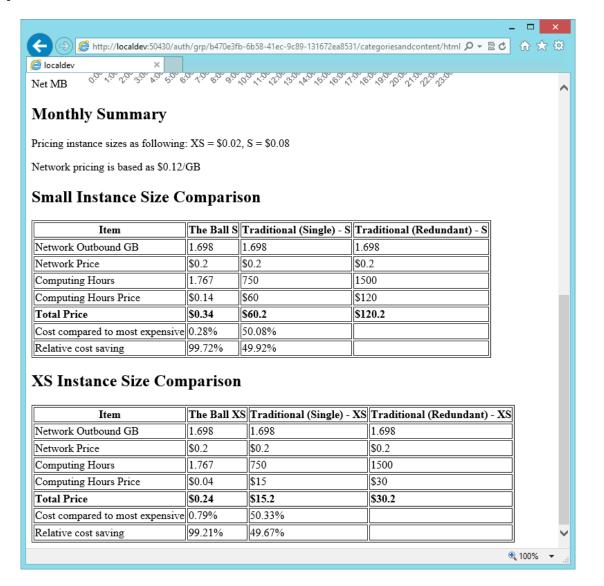
Simulation load scenario

- Load test content was decent sized website with 30 content pages
- Midnight: "visitor-simulator"
 1000 times home page load + few article clicks
- Running the batch operations on "remote browser" around 7:00 to calculate usage reports
- 9 10: 300 edits of articles
- 10-11: "visitor simulator" for 500 times





Comparison the Ball vs traditional



Demo - time