



Managing the Performance of Clouds and Cloud Based Applications

Alan Clark, Telchemy

alan.d.clark@telchemy.com

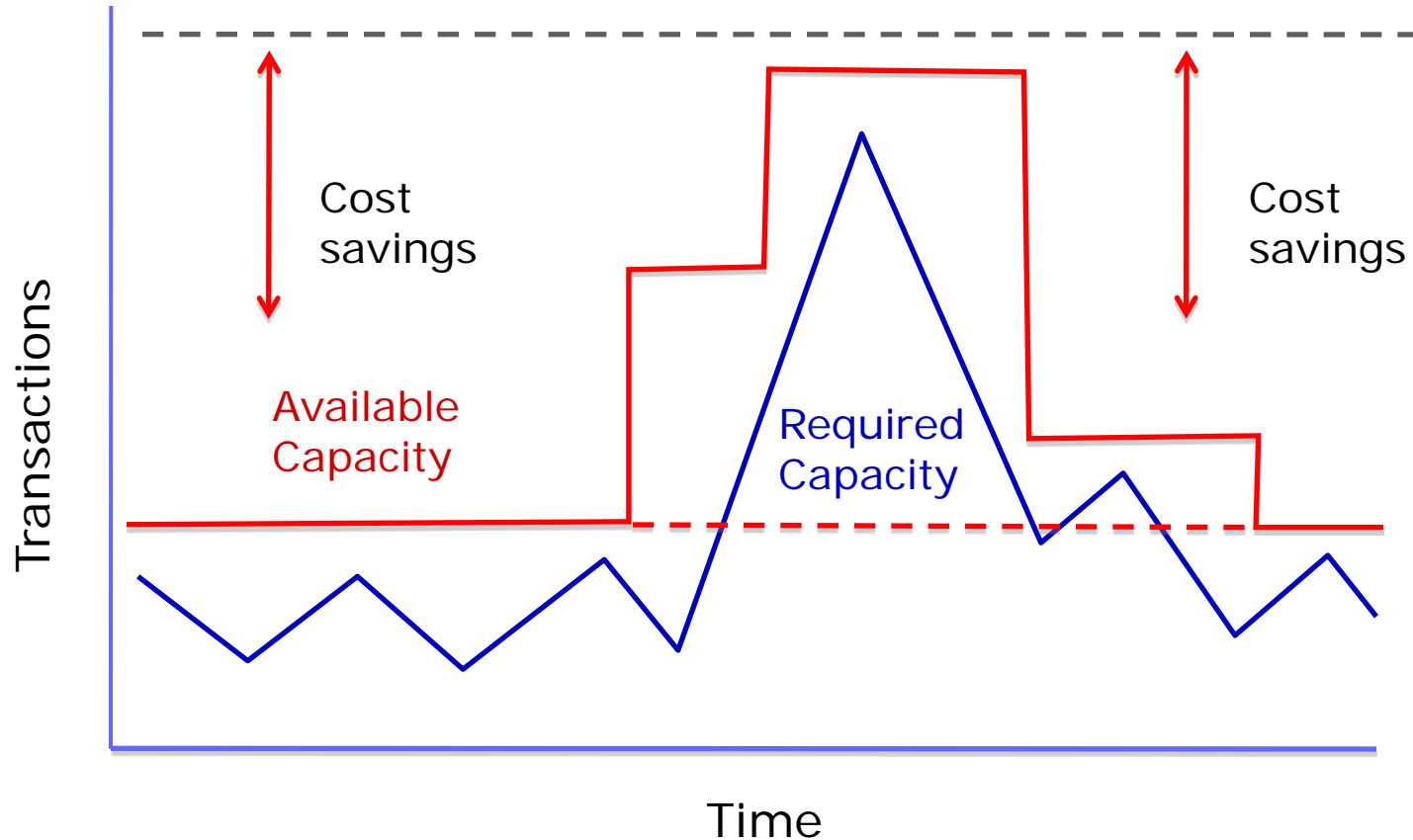
IEEE CQR 2012

Outline

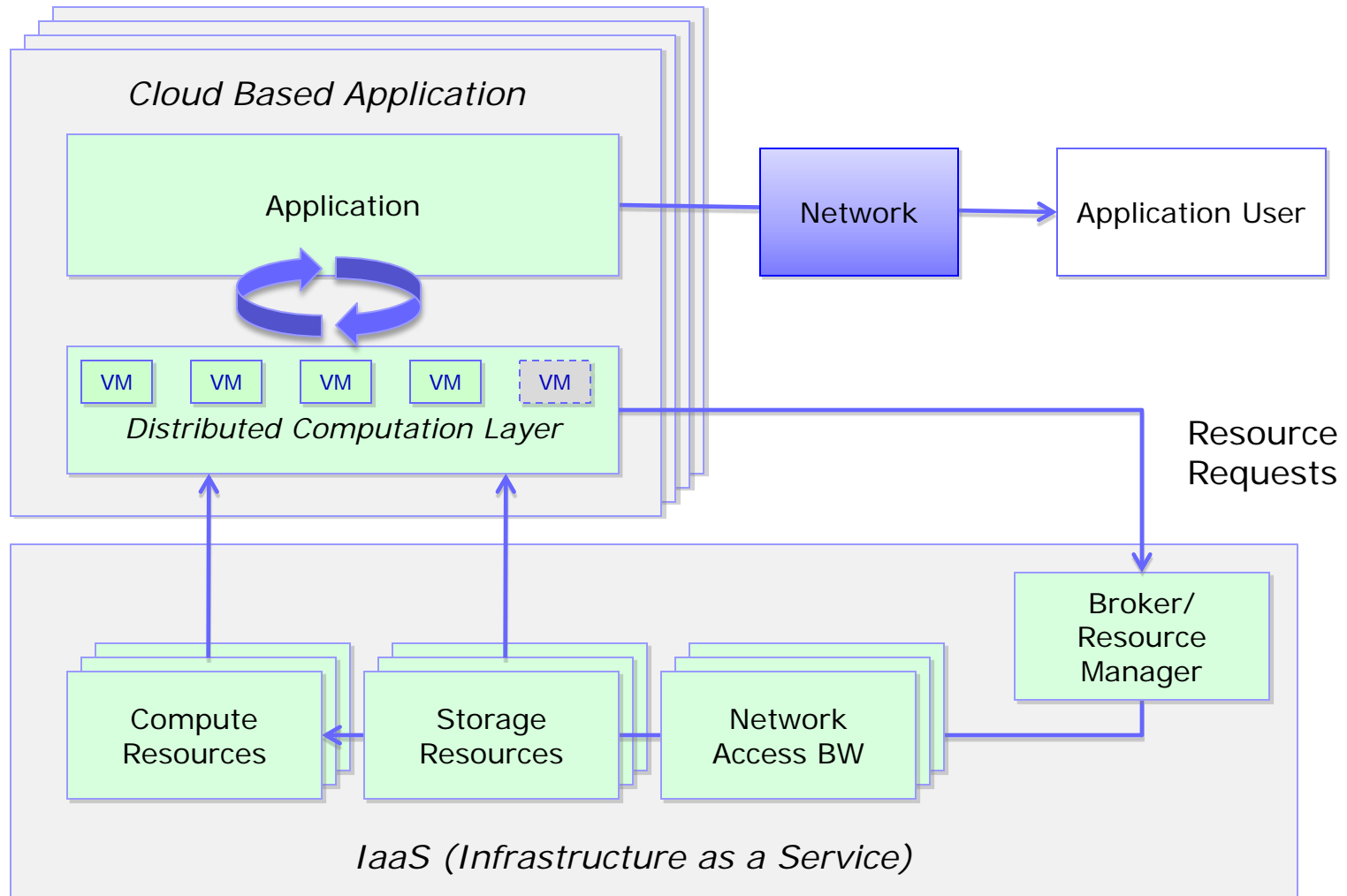
- Model for Cloud based Applications
- Cloud Application Performance
- Typical Cloud SLA
- What Cloud Infrastructure and Cloud Application providers need to measure

Why run applications in the Cloud?

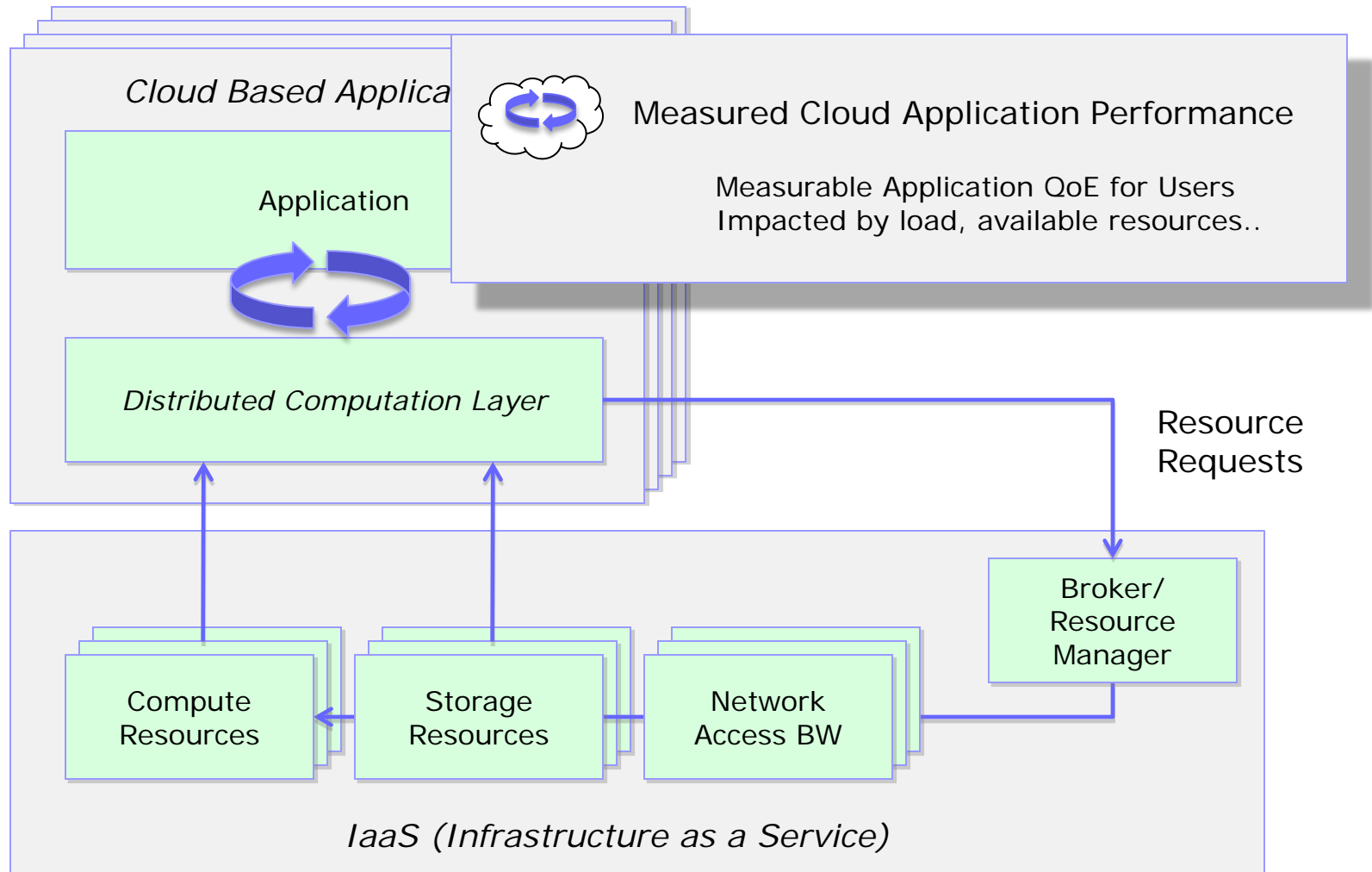
Flexible/ Elastic Resources – Computing on Demand



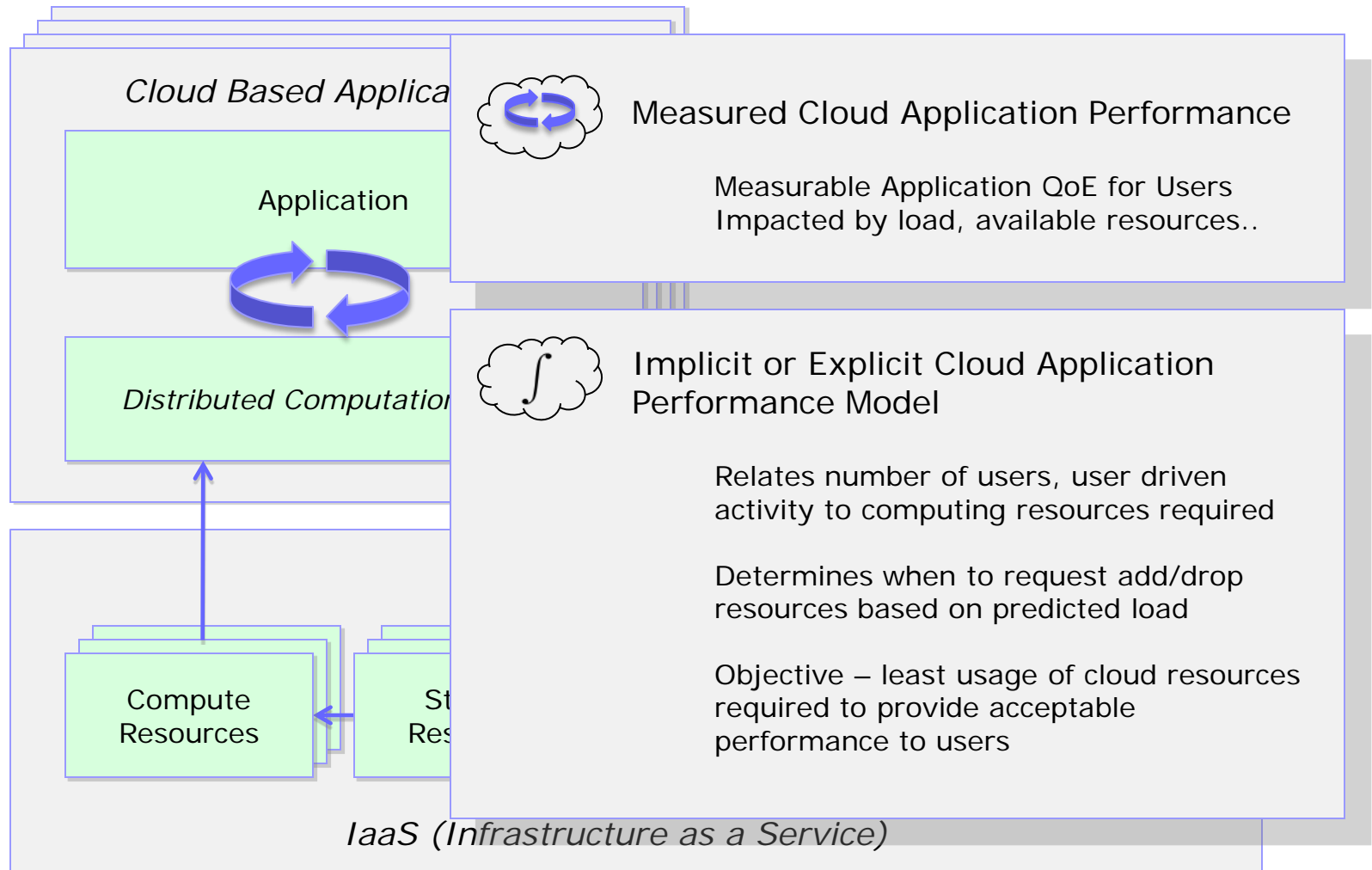
Model of a Cloud Based Application



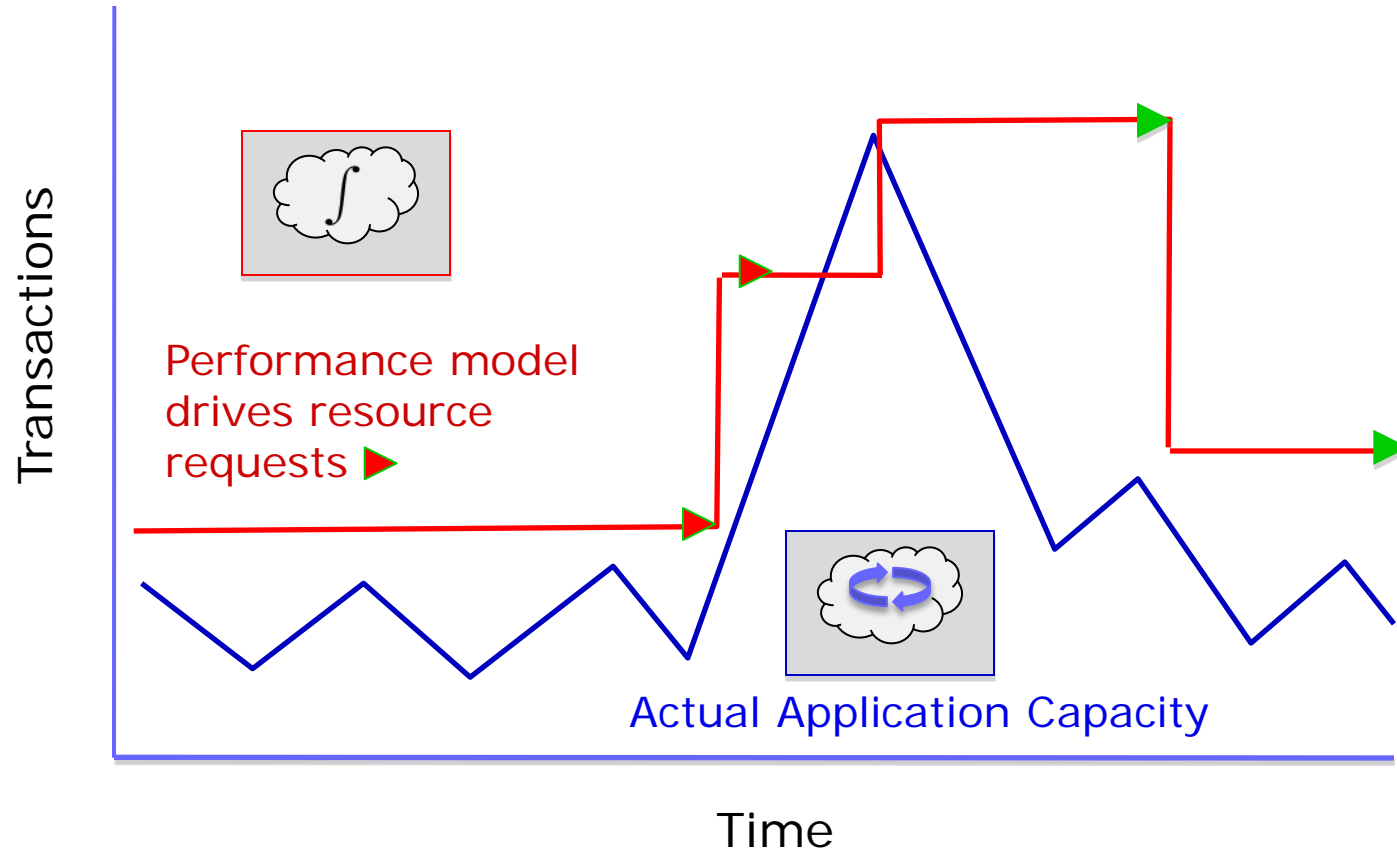
Cloud Model



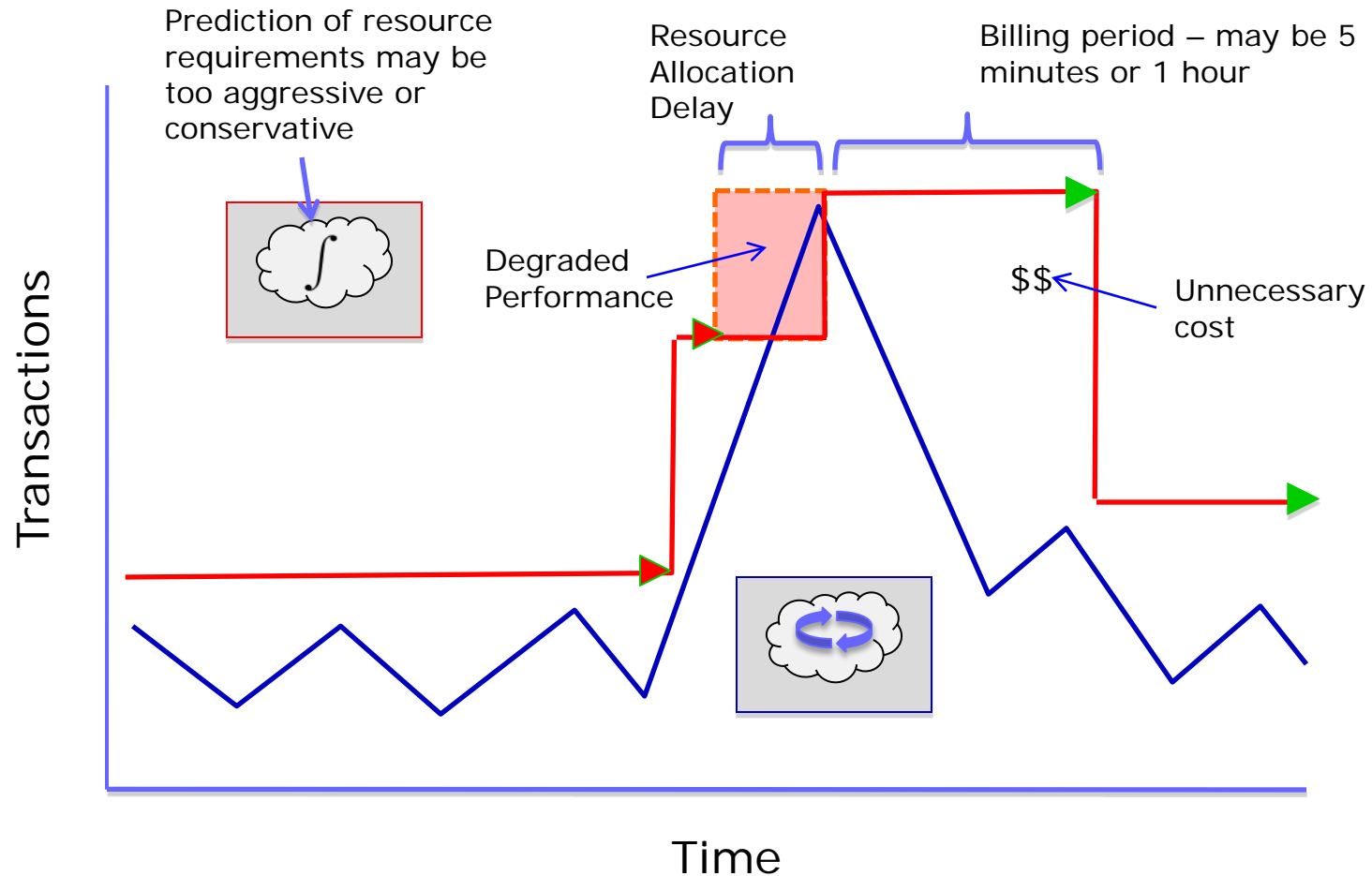
Cloud Model



Flexing Resources



Flexing Resources – potential issues



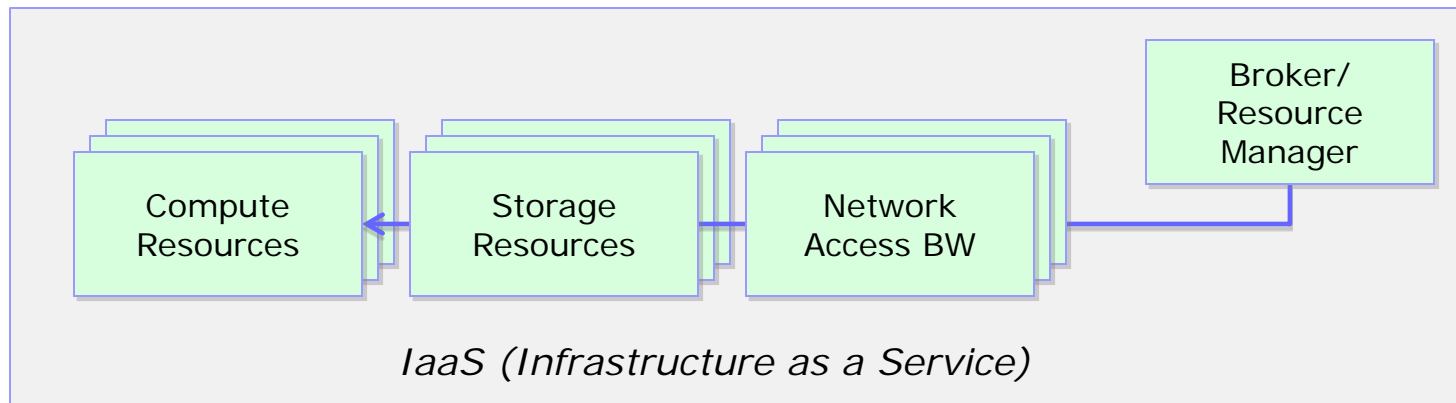
Cloud Infrastructure – typical SLA

Typical SLA – Availability

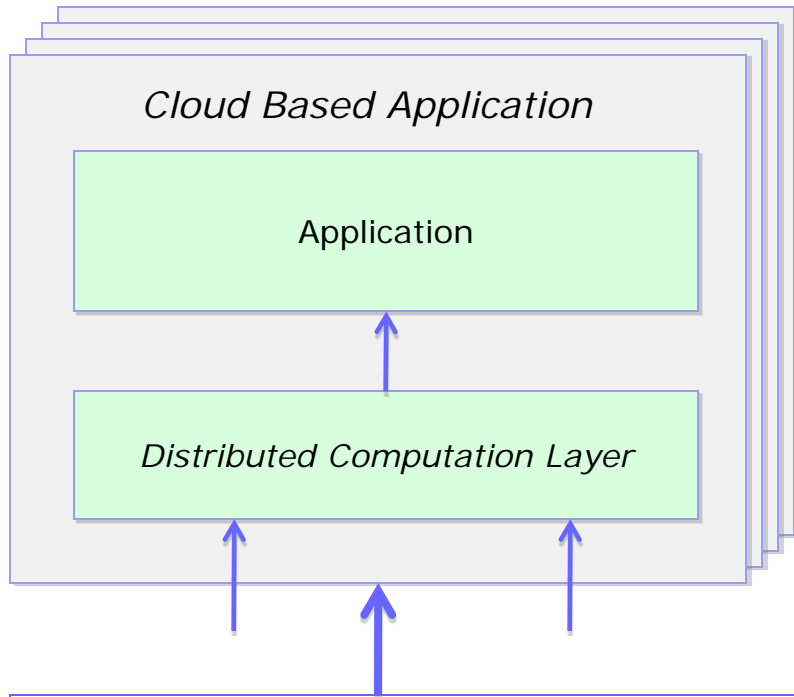
Amazon EC2 SLA - > 99.5% availability over 365 days
Verizon Terremark – service credit for each 15 minute outage

Typical performance metrics (not part of SLA)

CPU capacity allocated
Percentage usage of CPU capacity



What should a Cloud SLA guarantee?



Cloud SLA

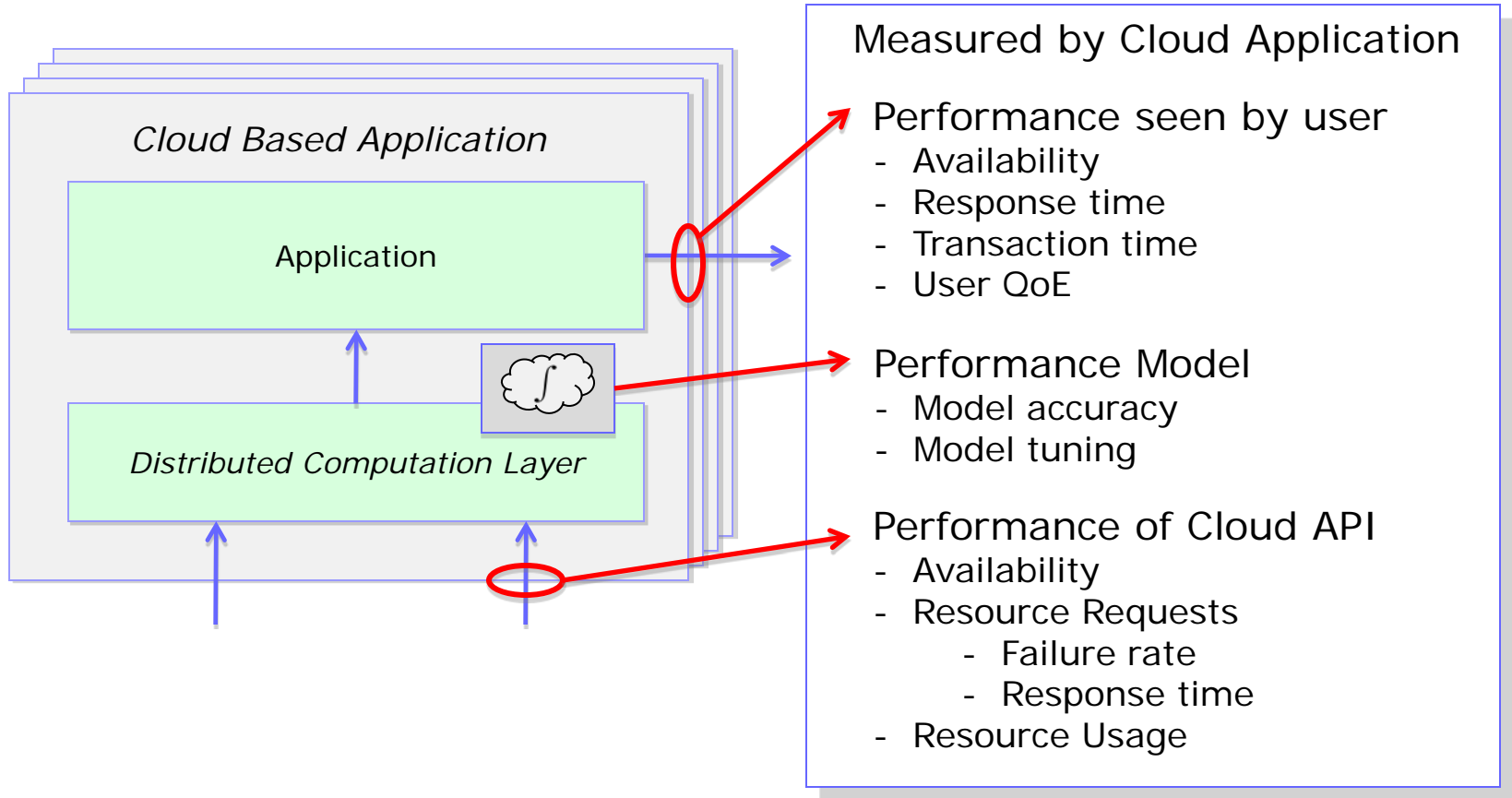
- Service Availability (Short/ Long Term)
- Dynamic Resource Availability
- Resource Allocation Response Time
- Resource Allocation Accuracy

Elastic Computing service – should have SLA that supports Elastic Computing

- Do resources get allocated
- Delay to provide resources
- Are correct resources being provided

Availability guarantees should be short term – e.g. minutes of outage per hour

What should the Application Provider measure?



Summary

- Model for Cloud based Applications
- Cloud Application Performance
- Typical Cloud SLA
- What Cloud Infrastructure and Cloud Application providers need to measure
- Related information
 - NIST Cloud Computing Definition
 - DMTF Cloud Management WG
 - ATIS Cloud Services Forum
 - ITU-T Cloud Focus Group
 - Amazon CloudWatch
 - OpenStack

About Telchemy

- Leading provider of technology for real time analysis of VoIP, Videoconferencing, IPTV and Video Streaming services
- Pioneered distributed performance management model for real time services, using endpoint based agents. Over 100 million agents deployed in IP phones, media gateways, CPE routers, MTA's, set top boxes, soft clients.....
- Performance Management solutions for IP Centrex/ Hosted PBX, Residential triple play, Mobile, Enterprise, Government