Our Sponsors

Scaling Analysis Services in the Cloud

by Gerhard Brückl

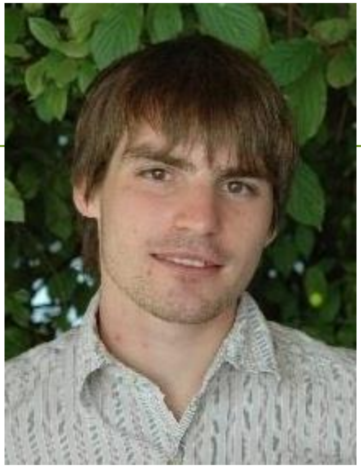
PASS

gerhard@gbrueckl.at
blog.gbrueckl.at

PASS SQL saturday

#280 | VIENNA 2014

About me



Gerhard Brückl

Working with Microsoft BI since 2006
Windows Azure / Cloud since 2013

focused on Analytics and Reporting

Blog: blog.gbrueckl.at
eMail: gerhard@gbrueckl.at

Agenda

- Why do we need to scale?
- How can we scale
  - Scale Up
  - Scale Out
- **Windows Azure**

Why do we need to scale?

- Growing amount of data
- Growing amount of users
- Actuality of data
- Complexity of data loads
- Latency (remote locations)

Things we want to achieve

- Better Query Performance
  - Single User
  - Concurrent User
- Faster availability of data
- High Availability
- Easy maintenance
- Flexible resource usage / peak times
- Better Processing Performance

How can we scale?

Scale Up

- Increase resources of current machine

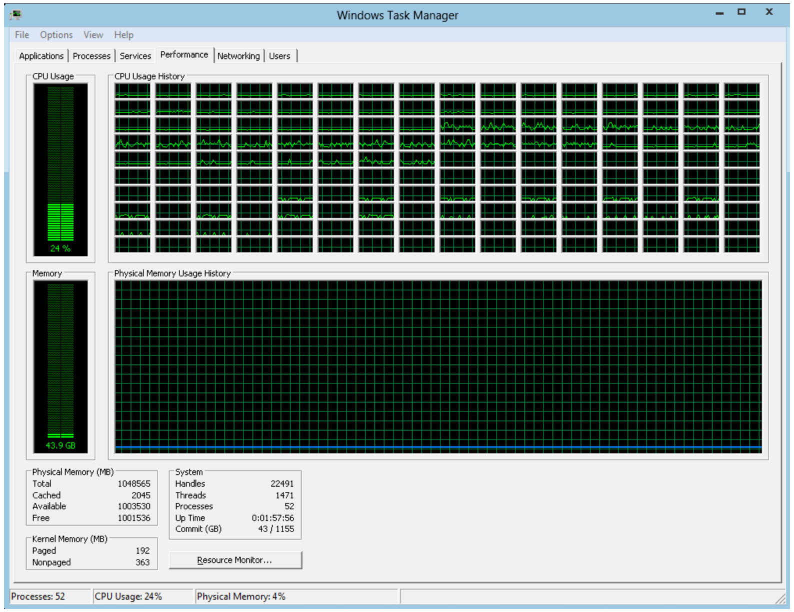
Scale Out

- Add further machines
- Create “Farm”

Scale Up

Increase

- CPU (clock rate vs. #cores)
- Memory
- IO
- (Network)



Scale Up

Pros

- Solves most issues
- Straight forward / simple
- No change in architecture

Cons

- Can only Scale Up to a certain point
- Expensive in terms of high-end hardware

Scale Up

CPU

- Parallel Processing / Querying
- Concurrent users
- (faster calculations)

Memory

- Caching – relieve IO / CPU

I/O

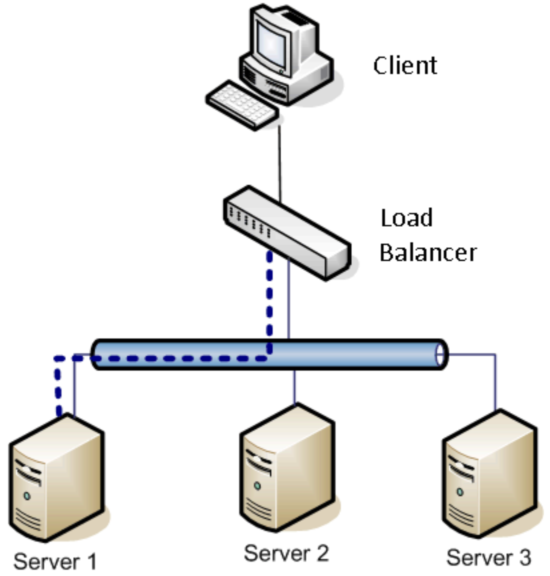
- Faster initial Querying / Processing

Scale Up

	Increase CPU	Increase Memory	Increase IO
Cube size (in general)	+	+	+
Query Performance (single user)	+	+	+
Query Performance (multiple users)	+	+	~
Processing Performance	+	~	+
Actuality of data	+	~	+
High Availability	~	~	~

Scale Out

- Increase number of machines
- Distribute work load
- Handle peaks



Scale Out

Pros

- "unlimited" scalability
- Flexibility
  - easy to extend / shrink farm
- Can be leveraged during data loads

Cons

- Change in Architecture
- Harder to maintain
- "Only" solves concurrency issues
- No improvements for single queries

Scale Out

	Increase Number of Machines
Cube size (in general)	~
Query Performance (single user)	~
Query Performance (multiple users)	+
Processing Performance	+
Actuality of data	+
High Availability	+

Windows Azure and Analysis Services

- Software as a Service (SaaS)
  - e.g. Office 365, SharePoint Online, …
- Platform as a Service (PaaS)
  - e.g. SQL Azure Databases, …
- Infrastructure as a Service (IaaS)
  - e.g. hosted virtual machine

Windows Azure IaaS

- Hosted Infrastructure
- Virtual Machines
- Virtual Storage
- Virtual Network

Windows Azure
Basic Considerations and setup

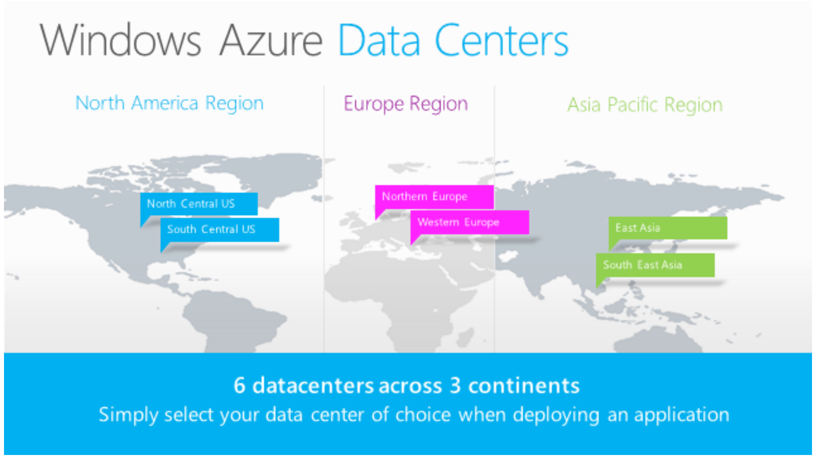
Choose a Region

Create an Affinity Group

- Per Region
- Per Subscription

Create Network

- Per Affinity Group



Windows Azure Virtual Network

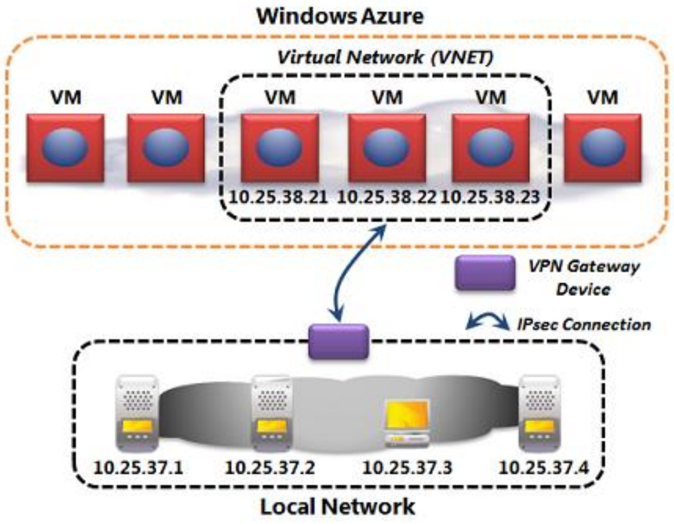
NETWORKS
1

Per Region / Affinity Group

Place VMs in same network

Integrate on-site services

- Point-to-Site Connectivity
- Site-to-Site Connectivity



Windows Azure Storage Account

Per Region / <u>Affinity Group</u>

Blobs / Tables / Queues

- VHDs

<u>Locally Redundant</u>

Geo Redundant

Geo Redundant (Read-Only)

Windows Azure Virtual Machine

Per Region / Affinity Group / Virtual Network

Created from Gallery

- Public Templates
- Private Templates

Different Sizes → Scale Up

Availability Sets → Scale Out

Windows Azure Scale Up

VIRTUAL MACHINES
1

Size	# of Cores	Memory	# of Disks	Price / Month *
Extra small	1 (shared)	768	1	315 €
Small	1	1,750	2	354 €
Medium	2	3,500	4	404 €
Large	4	7,000	8	504 €
Extra Large	8	14,000	16	1,008 €
A5	2	14,000	4	526 €
A6	4	28,000	8	747 €
A7	8	56,000	16	1,495 €
(A8)	8	56,000	? / SSD	?
(A9)	16	112,000	? / SSD	?

*) SQL Server **SE** + Windows

Windows Azure
Scale Up

Simply change size

- Restart required

Limited Capacities

- CPU
- Memory
- IO

Use dedicated machines!

Windows Azure
Scale Up – Analysis Services

Things to consider:

- Adjust I/O sub-system
  - Add VHDs / rebuild Storage Pool
- Adjust memory limits
  - Absolute settings
- Adjust CPUs
  - Thread settings
  - Group Affinity

Windows Azure
Scale Up – Analysis Services

For frequent Scale Ups / Scale Downs

- Use default settings

- Use only relative settings
  - Memory settings in %
  - Thread settings 0 or <0

Windows Azure
Scale Up – Analysis Services

VIRTUAL MACHINES
1

On-Premise vs. Cloud

CPU x4

Memory x2

CPU	On-Premise	Cloud "A7"
NUMA Nodes	2	2
Logical Cores	36	8

Memory	On-Premise	Cloud "A7"
Physical Memory	512 GB	56 GB
SSAS / TotalMemory[illegible]	100 GB	47 GB (85%)

IO	On-Premise	Cloud "A7"
Disks (SSAS only)	2	2
RAID-Set	RAID-0	RAID-0

Windows Azure
Scale Up – Analysis Services

Processing Performance:

Test SSAS-DB: 26 GB

Test SQL-DB: 106 GB

	On-Premise	Cloud “A7”
Parallel Processing	00:13:15	01:25:41
Serial Processing	01:36:05	05:17:28

Windows Azure
Scale Up – Analysis Services

VIRTUAL MACHINES
1

Query Performance:

7 Queries, +1 User/min, 60 mins

Response time in Seconds	On-Premise	Cloud "A7"
Single Query1	3.25	15.30
Single Query2	51.60	86.30

Finished Tests	On-Premise	Cloud "A7"
After 20 minutes/users	347	166
After 40 minutes/users	871	346
After 60 minutes/users	1,356	480 (*)

Windows Azure
Scale Up – Analysis Services

Scales with Number of CPUs

Memory for concurrency

IO for bigger databases

Windows Azure
Scale Out

Easily create/add new VMs

- Images
- Script

Built-In Load Balancing

- Availability Sets
- Traffic Manager

Unused VMs create no costs

Windows Azure Availability Sets

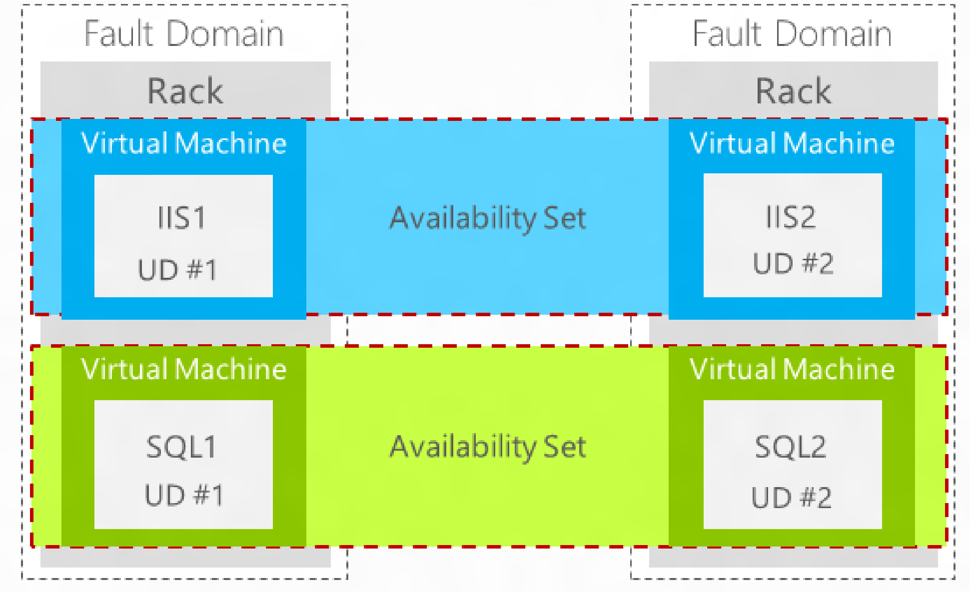
CLOUD SERVICES
5

Per Cloud Service

- Several VMs share same Public Port

Used for

- High Availability
  - Fault Domains
  - Update Domains
- Load Balancing



Windows Azure Traffic Manager

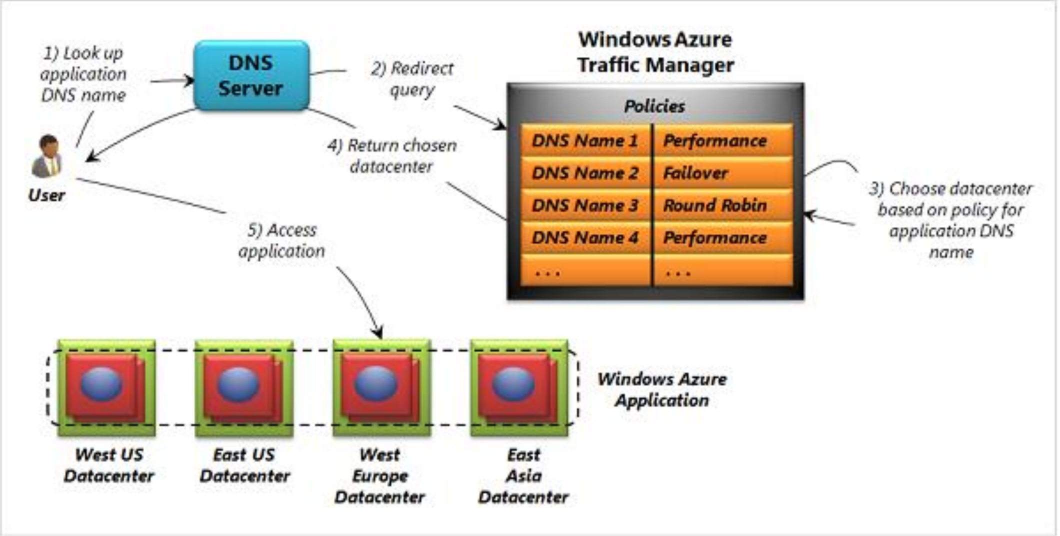
TRAFFIC MANAGER
1

Per Subscription

- Several Cloud Services share URL

Used For

- Availability
- Failover
- Performance



Windows Azure
Availability Sets and Traffic Manager

	Traffic Manager	Availability Set
Regions	Any	one
Round Robin	Yes	Yes
Performance	Yes	No
Failover	Yes	No
Routing-Level	DNS	TCP / UDP

Windows Azure Availability Sets

Auto-Scale Feature

- Scale by Metric
- Scale by Schedule

Only for Availability Sets

Also Third party tools

Windows Azure
Analysis Services Settings

Memory	Default Value	Suggested Value
LowMemoryLimit	65	85
TotalMemoryLimit	80	90
PreAllocate	0	85

OLAP/Process	Default Value	Suggested Value
AggregationMemoryLimitMin	10	5
AggregationMemoryLimitMax	80	10

Windows Azure
Scaling I/O for Analysis Services

Affinity Groups

Storage Account

- Locally Redundant
- Geo Redundant
- Geo Redundant (Read-Only)

Windows Azure
Scaling I/O for Analysis Services

Limits Storage Account (Locally Redundant)

- 10 Gb/s in / 15 Gb/s out
- 20,000 Transaction/s

Blobs and VHDs

- 500 IO/s
- 60 MB/s

Windows Azure
Scaling I/O for Analysis Services

Windows Server 2012 Storage Pools

- Abstraction Layer
- (Software) RAID

GUI very buggy

→ User PowerShell instead!

RAID / Storage Pools vs. Single Drives

Windows Azure
Scaling I/O for Analysis Services

Theory – 16 Disks:

60 MB/s * 16 → 960 MB/s

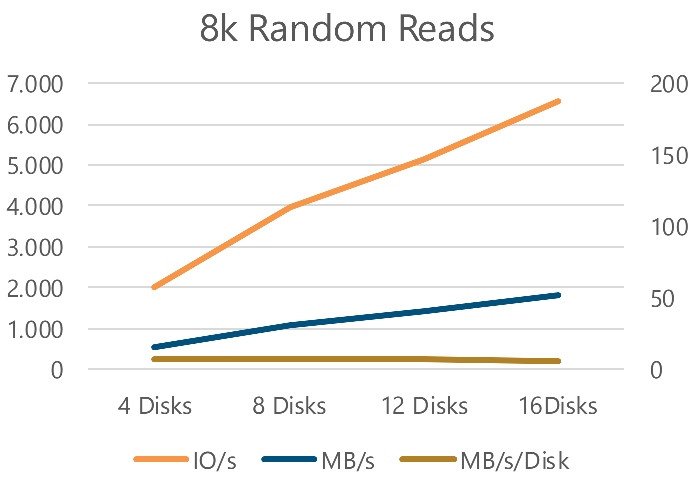
500 IO/s * 16 → 8,000 IO/S

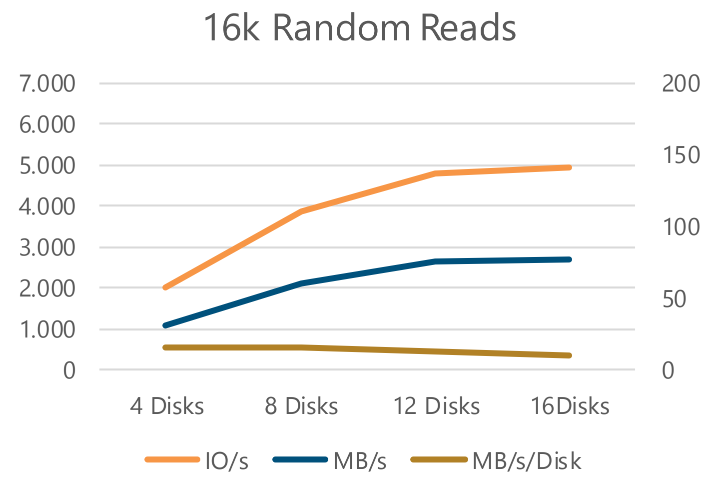
Reality – "A7" with 16 Disks:

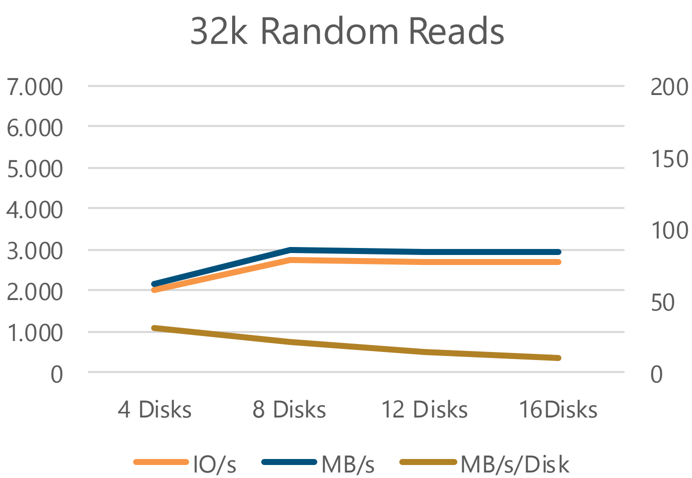
~110 MB/s

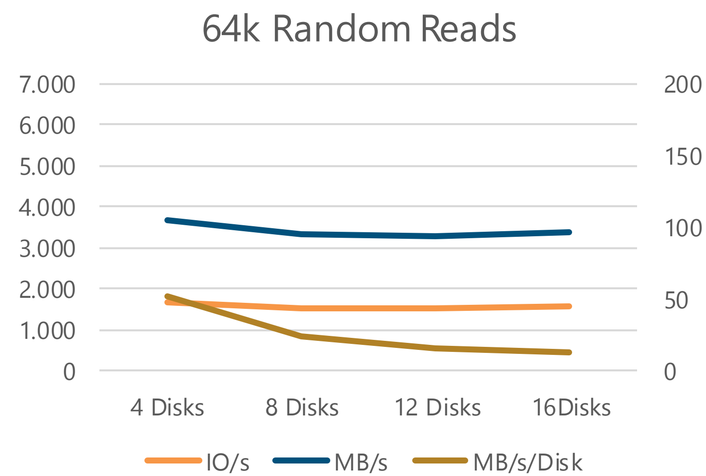
~6,500 IO/s

Windows Azure
Scaling I/O for Analysis Services

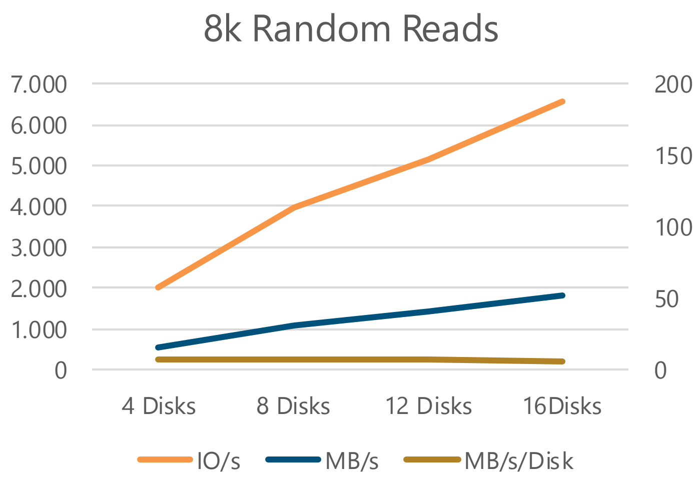


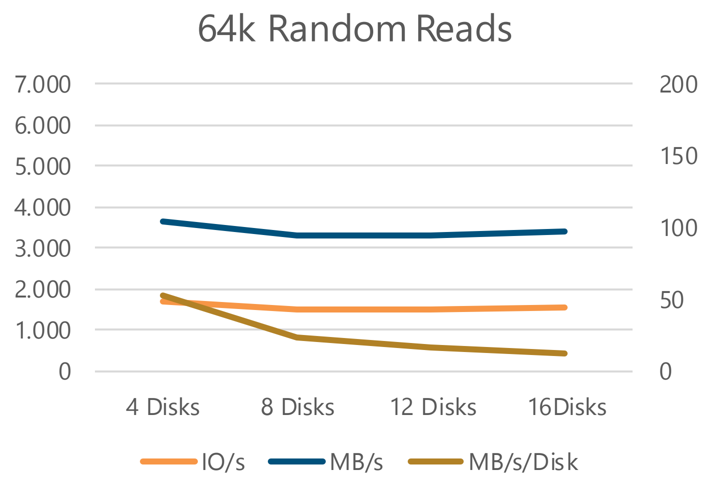






Windows Azure
Scaling I/O for Analysis Services





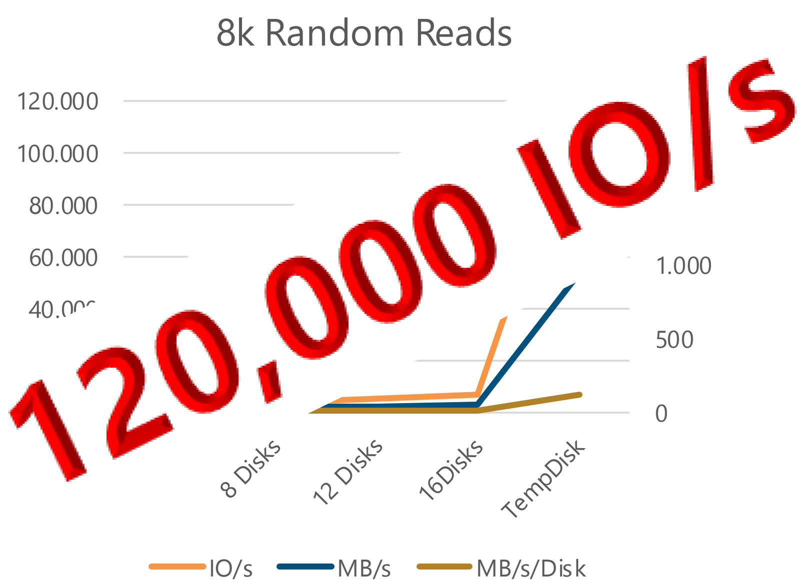
IO scales linear for small blocks

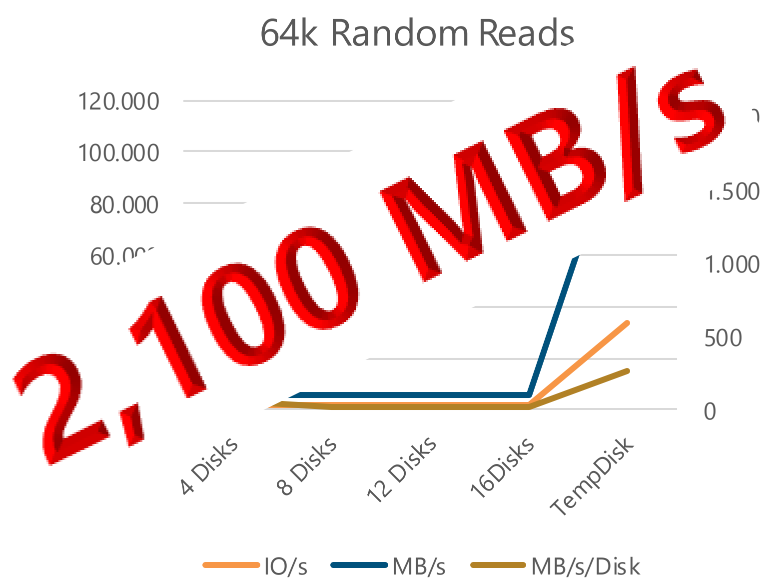
No improvements for bigger blocks

Windows Azure
Scaling I/O for Analysis Services

The special Temporary Storage (D:\)

- superfast but volatile storage





Windows Azure
Scaling I/O for Analysis Services

MSDN: "…performance is not guaranteed to be predictable"

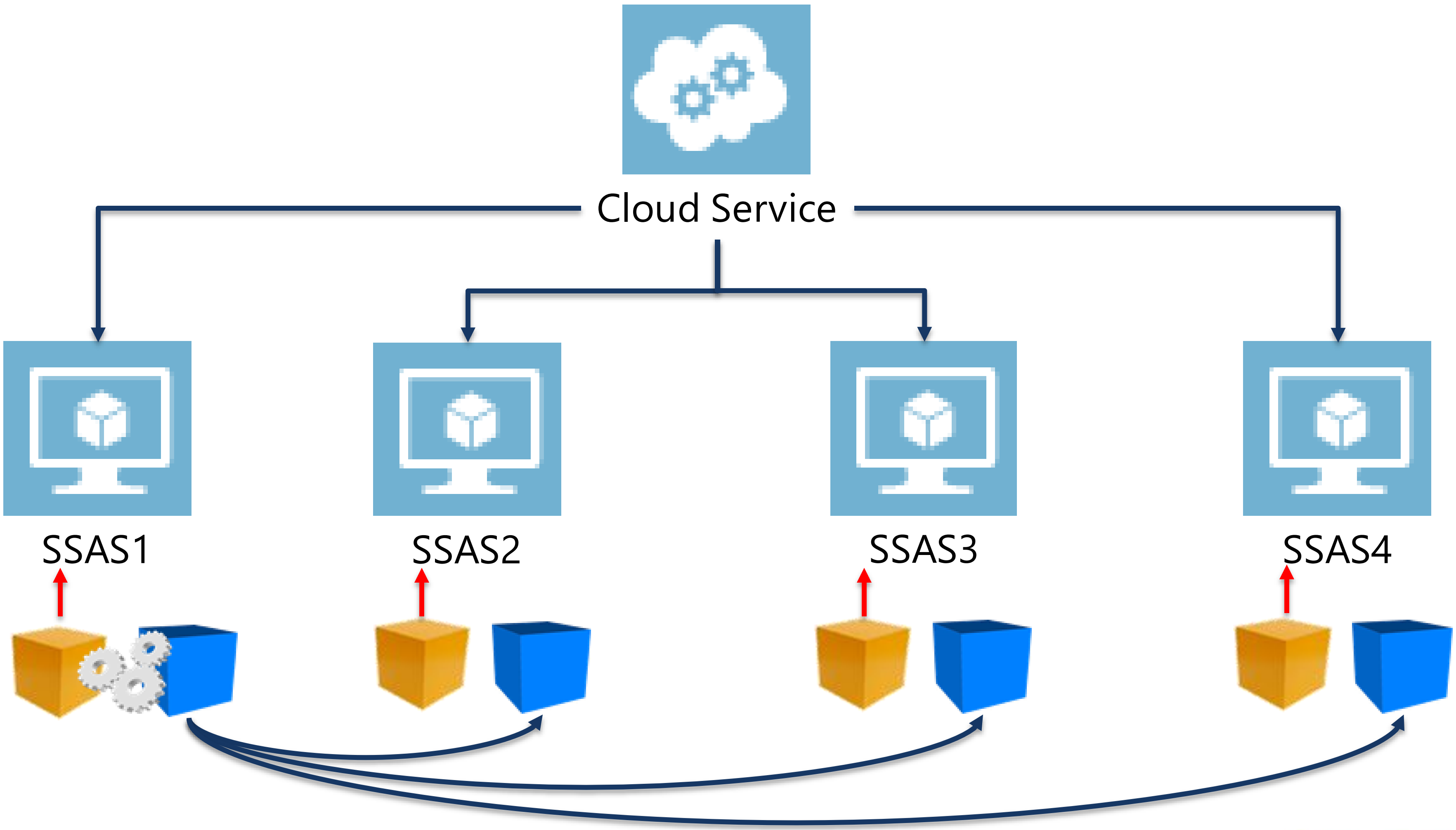
Deleted on resets, reboots, fail overs etc.

Scale-Outs with separate Query and Processing Servers

Windows Azure
Scaling I/O for Analysis Services

Distributing IO – Remote Partitions

Windows Azure
Processing Analysis Services



Conclusion & Findings
Scale Up

Easy to do

Only limited scalability

Watch out for:
- Disk / IO Setup
- absolute SSAS Settings

Conclusion & Findings
Scale Out

Almost everything can be automatized

- Create VMs
- Setup SSAS

Very flexible

- Shutdown unused VMs
- Auto-Scale Service

Conclusion & Findings
Scale Out

Use Availability Sets if possible

- Fault Domains
- Update Domains

Use Traffic Manager for

- Global Scale-Out
- Fail-Over

Conclusion & Findings
Scale Out

"Sticky Session" not supported!

Don't use "DirectReturn"

Possible issues with Session Objects

Calculated members/sets

- Excel Pivot Tables
- XL Cubed
- …

Writeback

Conclusion & Findings
IO

Use Storage Pools

Max # of Disks

Upcoming Changes

- A8 / A9 VMs
- Internal Changes

THANK YOU!