

veeAMON 2023

Configuration & Sizing Best Practices for Veeam Backup for Microsoft 365



Benedikt Daeumling
Business Solutions Architect
Veeam Software



David Tosoff
Senior Solutions Architect
Veeam Software

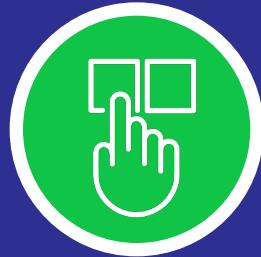


Agenda

- 
- 10:00 Introduction
 - 10:03 Visit the tailor
 - 10:05 Take your measurements
 - 10:15 Choose your materials
 - 10:25 Make the suit
 - 10:35 Final alterations
 - 10:40 Q&A
 - 10:45 Wrap up

Veeam Backup for Microsoft 365

Veeam® Backup for Microsoft 365 eliminates the risk of losing access to your Microsoft 365 data, including Exchange Online, SharePoint Online, OneDrive for Business and Microsoft Teams.



Infrastructure freedom

- Any deployment: on-premises, cloud, BaaS.
- Any storage: on-premises, cloud.
- Any change: add/change storage, location anytime.



Backup control

- Lowest RPO: as low as 5 mins.
- Tiered protection: different backup frequencies for different users.



Recovery flexibility

- Any recovery: broadest set of recovery options (50+ recovery options).
- Recovery out of M365: to on-premises, via email.



Scalable recovery

- Powerful search: search across users, many search filters.
- Bulk recovery: restore many users at once to effortlessly recover from disaster.

veeAMON 2023

Flexibility



Flexibility vs. Best practice

Many variables affect the needs, design choices and best practices in V365.

There is no one-size-fits-all when it comes to architecture, but there are some BPs.

Analogy: a tailored vs. off-the-rack suit. The OTR will do the job, but the tailored suit: 

Source	Server	Target
- Microsoft regions.	- On-premises or cloud.	- Repository on-premises or cloud.
- Different account MS policy.	- BaaS ?	- Block or object.
- Service usage.	- Physical or virtual ?	- Regulations (ISO, GDPR, etc.).
- Account size.	- Size of environment.	- Backup retention.

Let's visit
the tailor...



veeAMON 2023

What are you looking for?



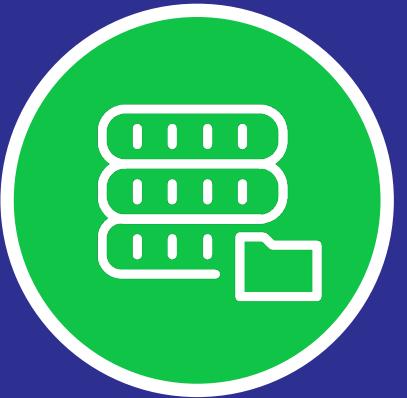
Key components



Backup proxy



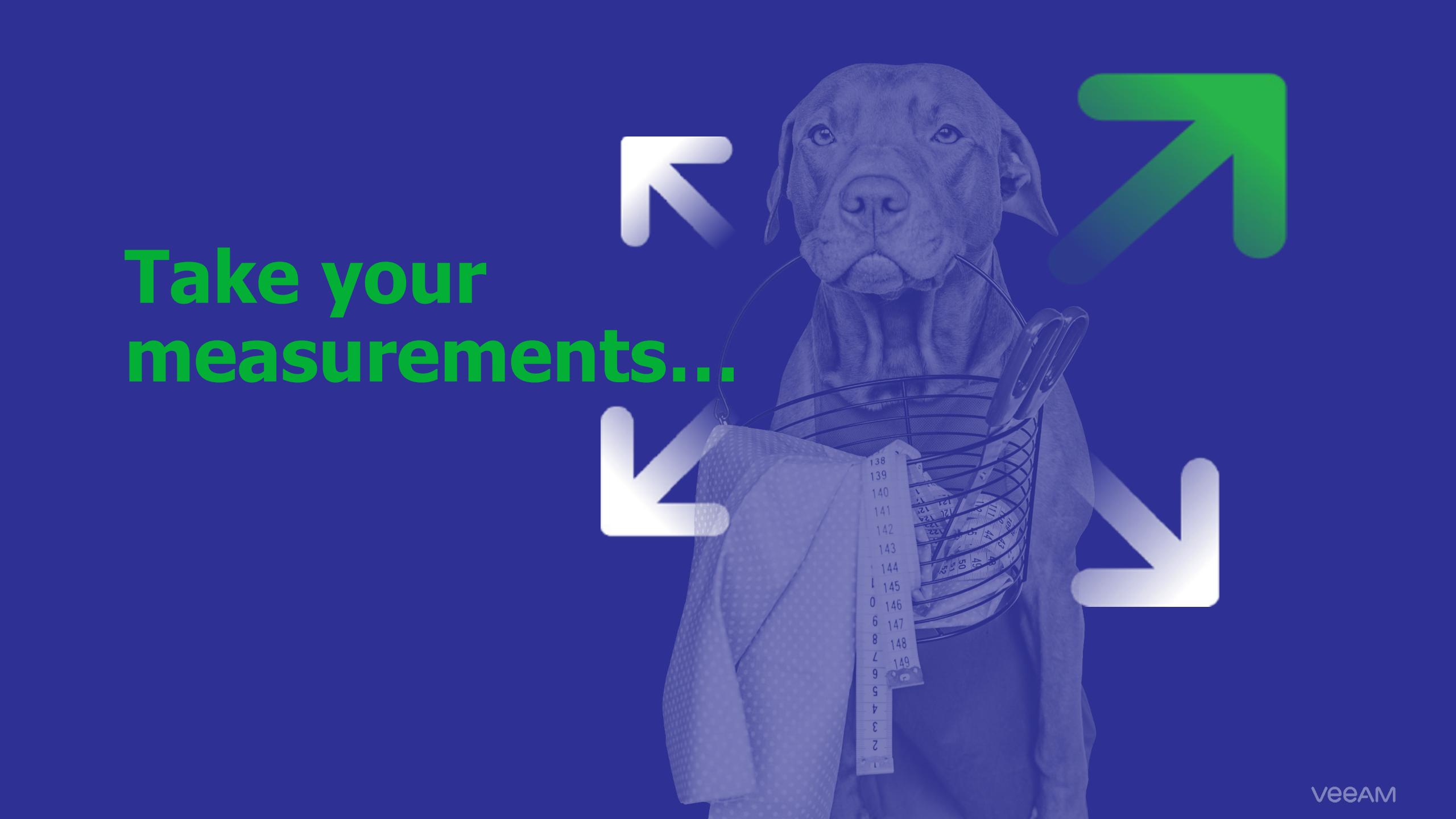
Backup server



Storage



Backup
repository



**Take your
measurements...**

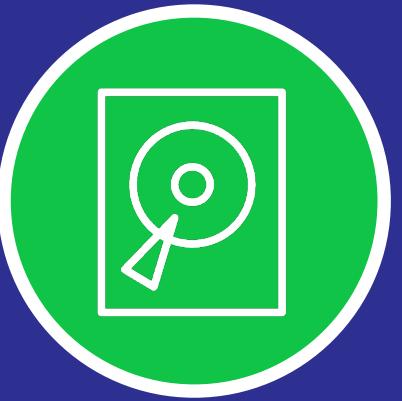
What is needed for a proper sizing?



Services to protect



Amount of objects



Retention requirements



Quantity of data
(and change rate)

What is needed for a proper sizing?



Services to protect



Amount of objects



Retention requirements



Quantity of data
(and change rate)



What are objects?

This is the main metric that affects proxy sizing...



Objects

1x object = 1x “thing” that we will back up the contents of...

The total number of objects managed determines our proxy(ies) sizing.

User object types:

- Mailbox.
- Archive Mailbox.
- OneDrive.
- Personal site.

Group object types:

- Group mailbox.
- Group site

Organization object types:

- SharePoint site
- Teams**

The total number of objects managed determines our proxy(ies) sizing.

**Count as 2 Objects

What is needed for a proper sizing?



Services to protect



Amount of objects



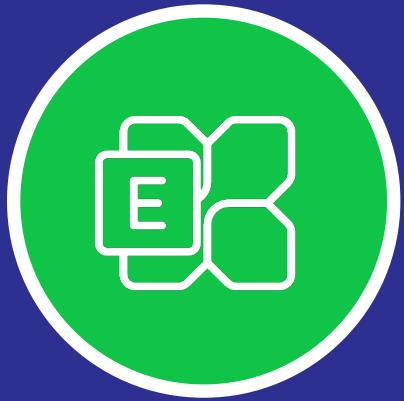
Retention requirements



Quantity of data
(and change rate)

Which services?

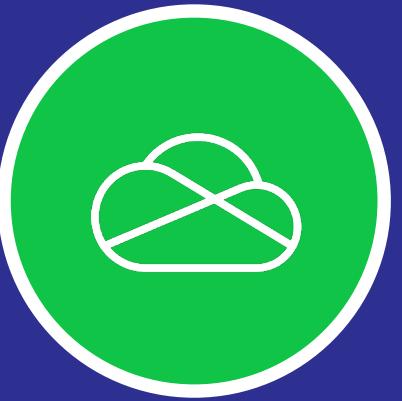
(And how many objects we have, **per service**)



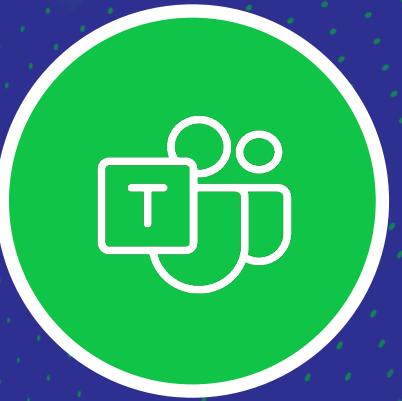
Exchange



SharePoint



OneDrive



Teams

Considerations

- Do you need to back up personal sites?
 - 90+% of organizations don't actually use (or often don't have a requirement to protect) these.
- Do you need to backup Teams chats?
 - Microsoft charges for this API (~\$0.75/1000 messages).
- Any compliance requirements you may need to adhere to?
- Would you like a French cuff or a barrel cuff?
 - You will need cufflinks to go along with that.

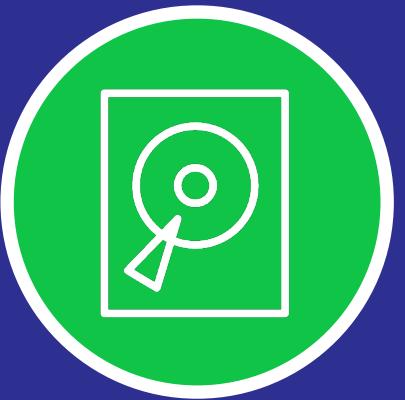
What is needed for a proper sizing?



Services to protect



Amount of objects



Retention requirements



Quantity of data
(and change rate)



What's my change
rate?

It's a little tricky, but...

Change rate

How much new or changed data is occurring in the M365 environment.

How do you measure this?

- “Easy” approach:
 - $(([\text{Current Data}] - [\text{X Days Ago Data}]) / [\text{X Days Ago Data}]) / [\text{X Days Ago}]$.
 - E.g. @ 10 days: $((200\text{TB} - 100\text{TB}) / 100\text{TB}) / 10 \text{ days} = (100\text{TB}/100\text{TB})/10 \text{ days} = 10\% \text{ per day}$.
- Steady-state approach: change rate per-day + filter to percentile, to even out major anomalies 
 - (e.g. one-time migrations may skew your change rate)
- Easiest approach: get-VB365Data script @
<https://github.com/VeeamHub/veeam-calculators>

What is needed for a proper sizing?



Services to protect



Amount of objects



Retention requirements



Quantity of data
(and change rate)

Retention

Organizational requirements?

Compliance requirements?

Where will the data be stored/retained?



Retention policy type:

- Item-based vs. “snapshot”-based.
 - Snapshot-based: age of the restore point/backup.
 - Item-based: age of the data/items themselves.
 - Cannot be changed once set.
 - Consider at outset of project. 
 - If in doubt, snapshot is likely what you want.

You're all measured up!



Services to
protect



Amount of
objects



Retention
requirements



Quantity of data
(and change rate)

Veeam Backup for 365 calculator

	Primary mailbox	Archive mailbox	OneDrive	SharePoint
Total database size, TB:	0 ▲ ▼	0 ▲ ▼	0 ▲ ▼	1 ▲ ▼
Weekly change rate, %:	0 ▲ ▼	0 ▲ ▼	0 ▲ ▼	0 ▲ ▼
Mailboxes:		Mailboxes:		Accounts: i
Quantity:	0 ▲ ▼	0 ▲ ▼	0 ▲ ▼	0 ▲ ▼
Retention policy i	Specify days ▼ 1 ▲ ▼	1 year ▼	1 year ▼	Specify days ▼ 1 ▲ ▼
Teams	10000 ▲ ▼			



Calculator secret decoder ring*

Assumes snapshot-based retention.

Each Teams is counted as two objects.

Local storage:

- Assumes no compression and adds 10% for working space.

Object storage:

- Assumes 50% compression.
- Assumes ~1.2-1.5MB average object size, depending on service.
- Does not factor for immutability API calls.

Servers:

- All-in-one: <20,000 objects** @ 8-core/32GB.
- Split backup server & proxy(ies): up to 20,000 per proxy @ 8-core/32GB.

BP

*Based on reverse-engineering

**20,000 aggressive for AiO



Choose your
materials...



The proxy choice

Location:

- On-premises – is copy job required (archiver/local-storage)?
- In-cloud – will egress be an issue?

Geographic:

- Compliance requirements?

Authentication:

- Domain-attached.
- Work group. 

Recommended Maximums

These maximums are based on Veeam QA testing and the recommended maximums for the specified config.

Type	Recommended Machine Config (8C/32GB)	Minimum Machine Config (4C/8G)
Proxies per VB365 controller	50	10
Objects per proxy ¹ (Refer)	20,000	2,000
Objects per VB365 installation ¹ (Refer)	1,000,000	20,000
Objects per organization ¹ (Refer)	400,000	20,000
Users per job	5,000	500
Users per proxy ² (Only)	5,000	500
Users per VB365 installation	250,000	2,500
Users per organization	100,000	2,500

That may not look good...

On-premises:

- Placement on-premises, if the WAN capability is limited
 - E.g. general WAN limitations.
[M365]→[WAN]→[Proxy]→[Storage]
 - E.g. hairpinning of traffic.
[M365]→[WAN]→[Proxy]→[WAN]→[Cloud Object Storage]

In-cloud:

- Burstable compute node types (e.g. Azure B-Series).
 - Errors, once burst credits are consumed.
- Placement in-cloud, if the target will be external object storage.
 - E.g. AWS EC2 w/ Wasabi storage.
 - It will work perfectly fine; however, cloud egress fees.



The repository choice

Type:

- Block (Local, SAN, SMB3* storage).
- Object.
 - Much higher compression.
 - Easier scale.
 - Supports encryption at-rest directly.
 - Supports backup copy job.
 - Supports immutability (via copy job).

Location:

- On-premises.
- In-cloud.

*SMB3 is Experimental Support

That may not look good either...

Block storage:

- Deduplication, compression nor encryption.
- SMB-based storage.
- ReFS w/ integrity streams.

Object storage:

- Spinning disk (HDD) for cache locations.
- S3-compatibles that do not fully/correctly implement the stack.
- Storage tiering/lifecycle.
- Versioning enabled (unless S3 w/ object lock).



(Paisley should be used in moderation)



Make the suit...





Make the suit...

Gathering the materials (prep)

Use the same OS version for all components.

BP

Use modern app-only auth.

BP

Prepare for Teams API, if required (see KB4322).

Have a global administrator ready, or manually create your application registration(s).

- Note: GA credentials are only used one time and are not retained.
- Least-privilege: evaluate whether you need *all* permissions.

BP

Set application impersonation for future exchange restores.

Ensure Teams license on registrant's account.

Pick a hostname for your backup server (forever!)

BP

Ensure all servers are on the same time zone.

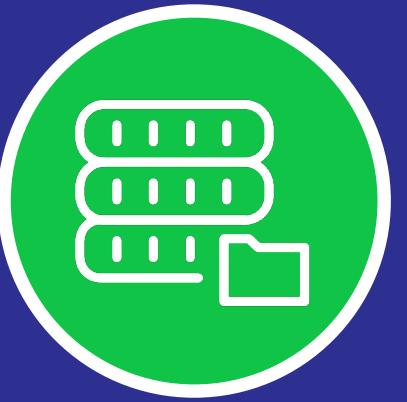
Key components



Backup server



Backup proxy



Backup repository



Storage

Making the jacket (server)

Backup server:

- 4-core & 8GB minimum.
- Use SSD for OS, logs and configuration (ProgramData) storage.

BP

Using REST API/self-service portal?

- + 4-core & 8GB.
- Split from Backup Server in large-scale deployments.

BP

Avoid co-installing w/ VBR.

BP



Making the vest (proxy)

Proxy server:

- 4-core & 8GB minimum.
- 8-core & 32GB recommended maximum.
- Use SSD for OS, logs and configuration (ProgramData) storage.

BP

BP

Start with the defaults.

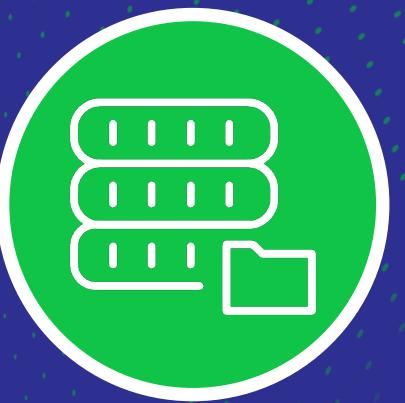
Making the pants (repository)

Repository:

- Decision-points for when to create an additional repo:



- Use SSD for object storage cache.
– Prepare for ~1% of source data size as cache.
- Retention policy type is configured once, at creation.
- Encryption & immutability are configured once, at creation (object storage only).
- Carefully consider immutability period.
– Data is locked for the full retention period set on the repository.
- Consider an abandonability strategy.



Making the pants (storage)

Object storage:

- Hot (S3) or cool (IA) or cold (archive/glacier) or frozen (deep archive)?
 - Primary backup: hot for shorter retention, cool for longer (~>45 days).
 - Backup copy: depends.
- How many buckets/containers? And storage accounts?
 - >1x per service type, per proxy, per location. More doesn't hurt. BP
- Use archiver appliance, if:
 - Proxies are on-premises or out-of-region (in the same cloud).
 - Backup copy jobs consuming proxy resources away from backup jobs.

Block storage:

- Always keep/add 10% free working space. BP
- Keep to a manageable size (<200TB).



Wearing the suit (jobs)

Design jobs based on:

- The target repository (→location, proxy, retention, ...).
- Split jobs up by service and reasonable user counts.
 - ~1000/job → goal being job completion.
- Schedule.

Object selection:

- Organization groups that already exist (e.g. distribution lists).
- Manually divide + catch-all job.
- Automate:
 - Azure AD dynamic membership (e.g. user objects).
 - Script to divide across jobs (e.g. Teams, SharePoint).
 - Consider GUID vs. Name, and -NotInJob switch.

BP



Teams backup

Teams uses SharePoint and OneDrive together.

- Data is approximately 70% SharePoint, 30% OneDrive.

Place Teams + SharePoint together in a job.

- Or at least using the same repository.
- Space savings with this method.

A second pair of pants/shirt (3-2-1)

...Just in case — you might spill something on them.

BP

How do you achieve 3-2-1?

- Backup copy job.
- Backup VB365 w/ VBR.



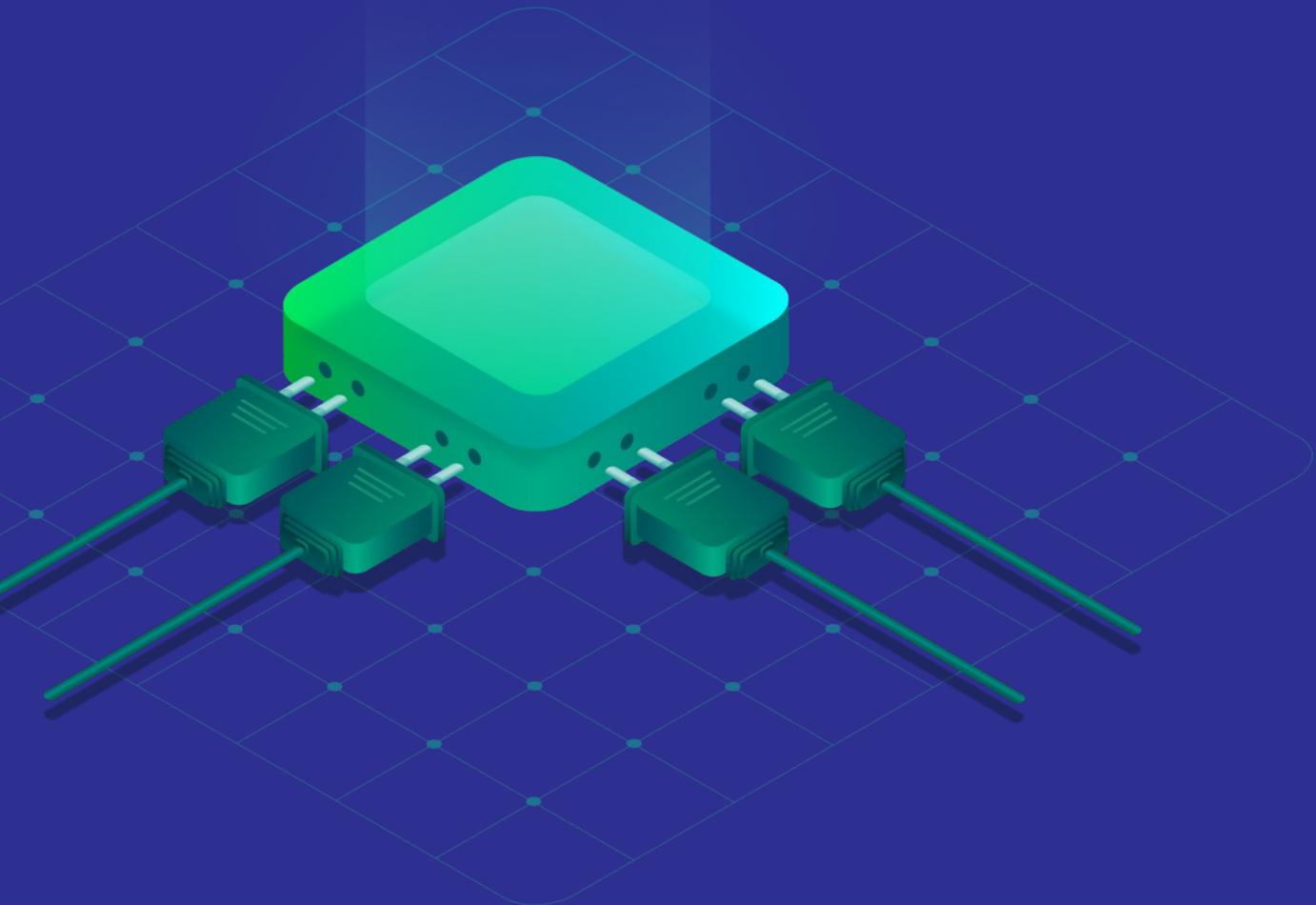
Isn't LRS enough? or, GRS makes a copy offsite for me...?

Considerations:

- No copy job for block storage.
- No immutability on primary backup.

Final
alterations...





Threads?

What are they?
How many do I need?



Threads

A thread is a connection to M365.

Handles data transfer for a subset of items in an object (e.g., folder, site/web, list).

The algorithm is a bit complex to describe.

The main takeaway: sometime less is more. Sometimes more is more.

- More threads \approx better CPU utilization, but more potential for throttle.

Default is 64.



Throttling?

Why are we being throttled?
How can we overcome it?

M365's throttling

You **will** experience it. It's part of Microsoft's service design, and it is normal.

Every request consumes resource units (RUs).

of RUs are per application registration; depend on your tenant size:

License count	0 - 1k	1k - 5k	5k - 15k	15k - 50k	50k+
App 1 minute	1,200	2,400	3,600	4,800	6,000
App daily	1,200,000	2,400,000	3,600,000	4,800,000	6,000,000

Source: <https://learn.microsoft.com/en-us/sharepoint/dev/general-development/how-to-avoid-getting-throttled-or-blocked-in-sharepoint-online>

There is also a per-tenant limit that can be reached (!!).

Requests that respond with throttle (429) still consume RUs.

More threads = more requests = more RU consumption = higher chance of throttle.

Am I being throttled?

Probably. Maybe. Yes? No...

Again, it's normal/by-design. And it may not actually be impactful.

Usually most impactful during initial sync (full backup).

Check the logs for 429 (too many requests) or 500/503 (internal server error).

- From proxy:

```
Get-ChildItem -Path C:\ProgramData\Veeam\Backup365\Logs -Recurse -Filter *.log | select-string "throttled [^o]"
```

- From VB365 server:

```
Get-ChildItem -Path C:\ProgramData\Veeam\Backup365\Logs -Recurse -Filter *.log | select-string "too many requests"
```

How can I overcome throttling?

First: determine if it's even an issue:

- Are backups completing in a reasonable amount of time?

Down-tune thread count on the proxy(ies) in question:

Add backup applications, with Veeam Customer Support's guidance.

- Recall that tenant throttle limit? We must be careful.
- Ensure to describe any other third-party applications that may be heavily using RUs.

Place exchange jobs on a different proxy.

Open a case with support for further performance tuning.



What else?

Any other alterations (tuning)
needed?

Further alterations

Performance:

- Be wary of using internet proxies (web proxy).
 - Fully supported, and configurable globally and/or per-proxy.
 - Can drastically affect performance if proxy overloaded.
- Avoid heavy use of exclusions in jobs; favour inclusive scoping over exclusive.
- If extended logging is no longer need it, turn it off.
- Set AV exclusions.

Security:

- Consider signed certificates for API & restore portal (self-signed by default).
- Avoid use of basic/legacy or legacy w/ modern authentication methods.

Helpers/scripts

Powershell:

- <https://github.com/VeeamHub/powershell>
- Any script that starts with **VBO-** or **VB365-**.
- Notably:
 - VB365-JobManager -- Create and distribute sharepoint sites across jobs by parameters.
 - VBO-SpJobsSiteID -- Create and distribute sharepoint sites across jobs by site ID.
 - VBO-CreateJobByOrg -- Create, exclude or split backup jobs by organization.

Get started with Postman/REST:

- <https://github.com/VeeamHub/veeam-postman>
- Check out the Veeam Backup for M365 subsection

Questions?





Thank you

veeAMON 2023