



# Wirtualizacja pamięci masowej i implementacja Software-defined Storage

Jarosław Biniek, Inżynier Sprzedaży  
[jaroslaw.biniek@suse.com](mailto:jaroslaw.biniek@suse.com)  
SUSE Polska





Mobile Data

# Eksplozja danych

Medical Data



IoT Data



175 ZB  
by 2025



Videos



Transactional Data



Emails

# Ille to jest zettabajt?

kilabajt =  $10^3$

...

terabajt =  $10^{12}$

petabajt =  $10^{15}$

eksabajt =  $10^{18}$

zettabajt =  $10^{21}$

- Odpowiada to 8 milionom lat materiału wideo w formacie UHD 8K



# Prawo Moore'a

Liczba tranzystorów podwaja się co ok. 24 miesiące

## Jak to się ma do danych?

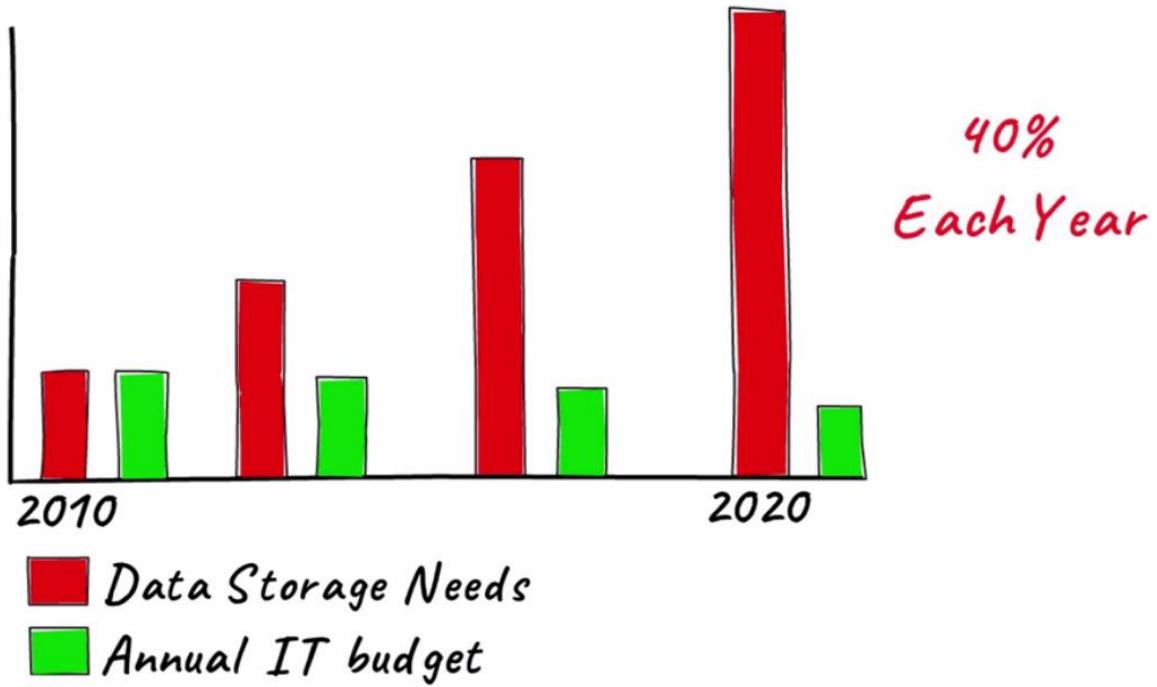
- Przyrost na podstawie raportów IDC
  - 244% (2011-2013)
  - 364% (2013-2016)
  - 1019% (2016-2025)



Personal Electronic Transactor

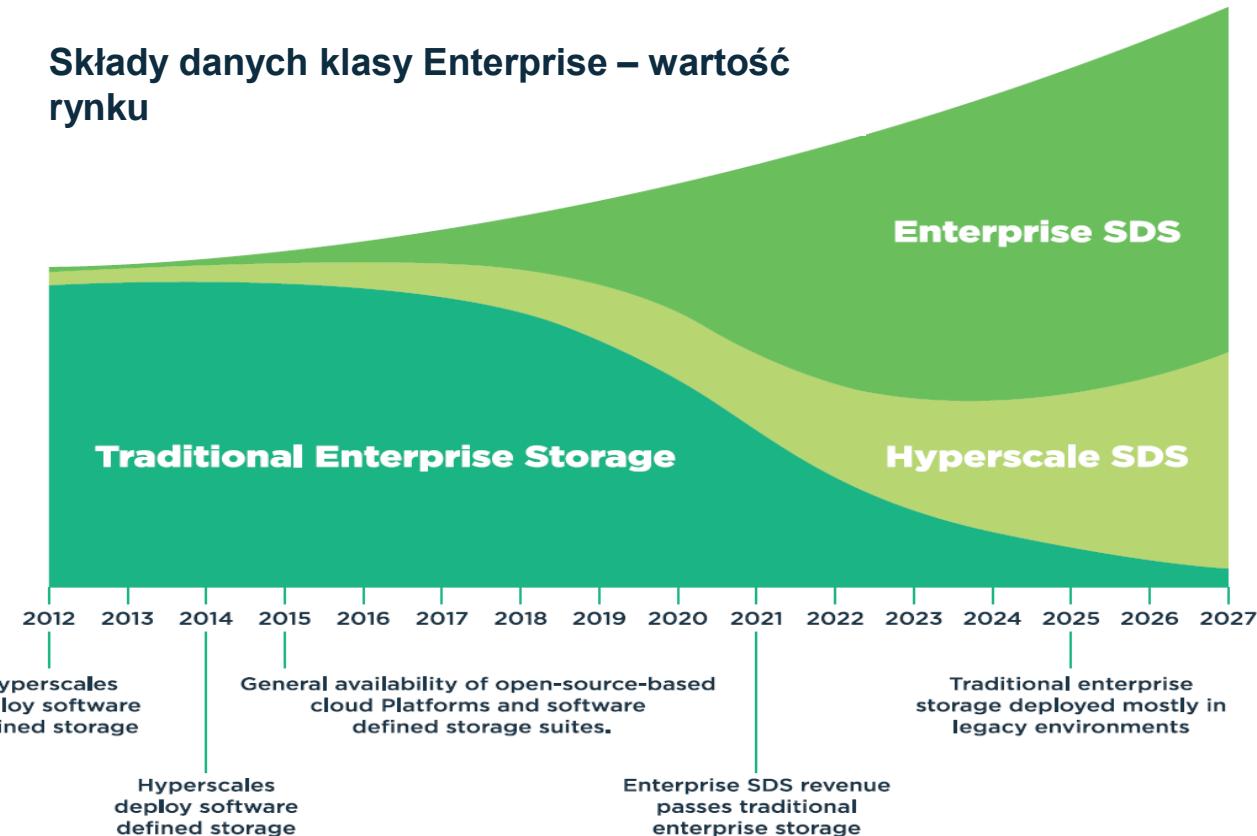


# Co to implikuje?



# Jak to wpłynie na rynek pamięci masowych?

Składy danych klasy Enterprise – wartość rynku



Opracowanie IDC i SUSE

# Czynniki ograniczające w tradycyjnych macierzach



Trudne do skalowania



Drogie



Trudne do zintegrowania w infrastrukturze definiowanej programowo

# Projekt Ceph

Inteligentne rozwiązanie z dziedziny software-defined storage, które pozwala przetransformować infrastrukturę do przechowywania danych, tak aby:



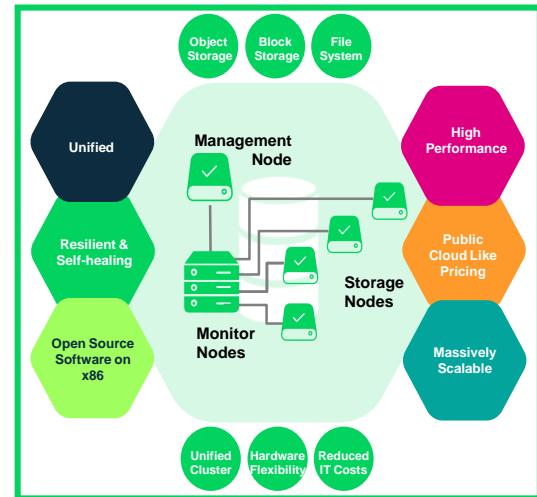
**Dostarczyć wysoko skalowalne i odporne**  
środowisko bez pojedynczego punktu awarii



**Obniżyć koszty** poprzez użycie zwykłych  
serwerów i dysków

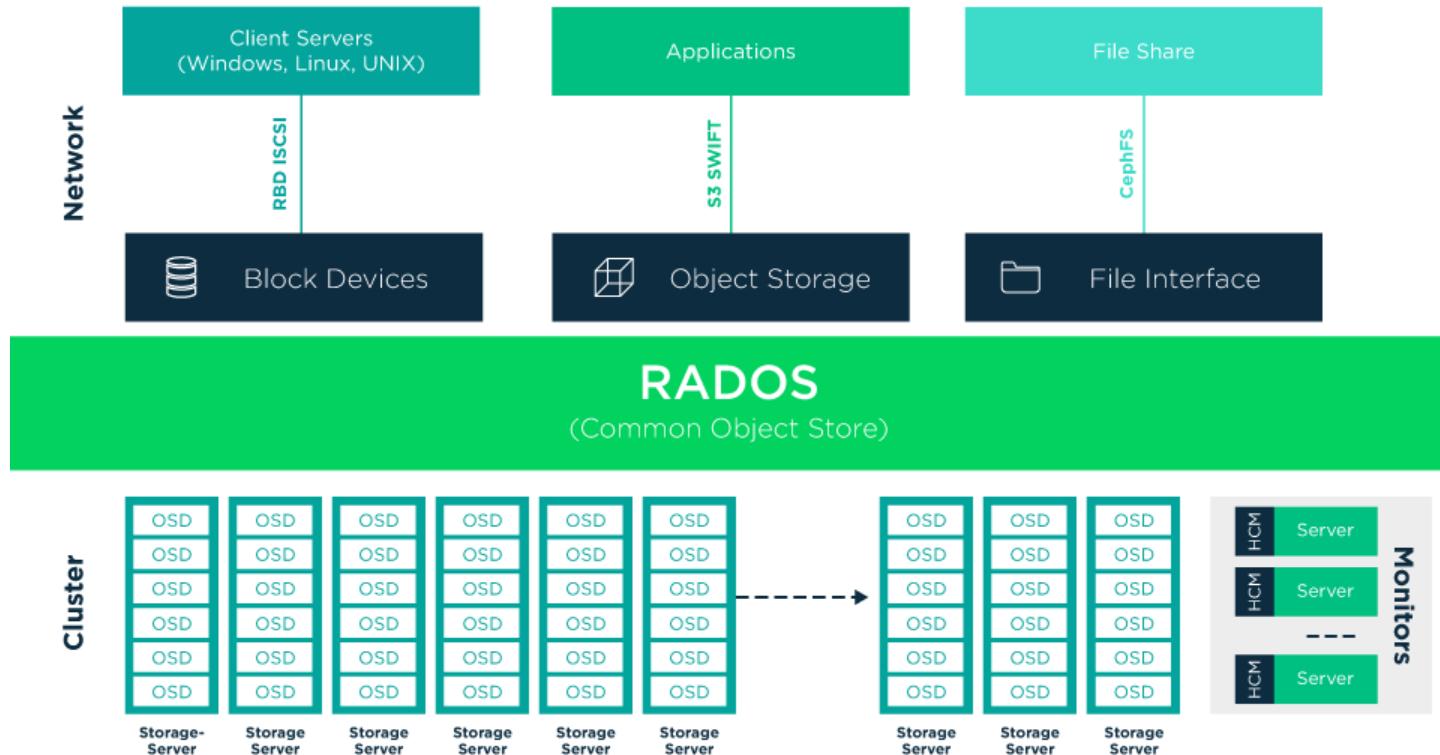


**Automatycznie optymalizować** i dodawać  
przestrzeni w miarę potrzeby, bez przestojów



# Napędzany technologią Ceph

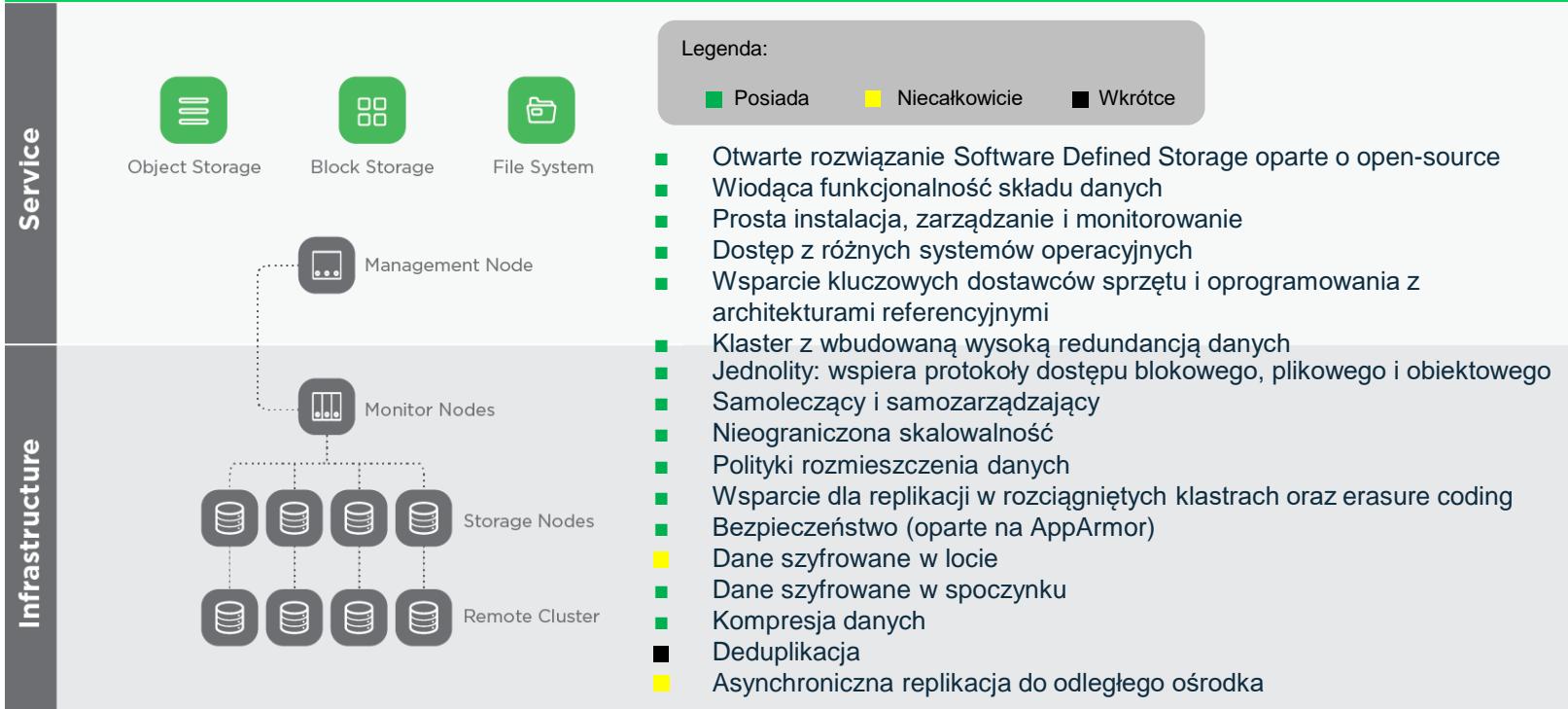
## Architektura SUSE Enterprise Storage



# Skład danych klasy Enterprise

## Kluczowe cechy i funkcje

### SUSE Enterprise Storage 6



# Wspierane protokoły

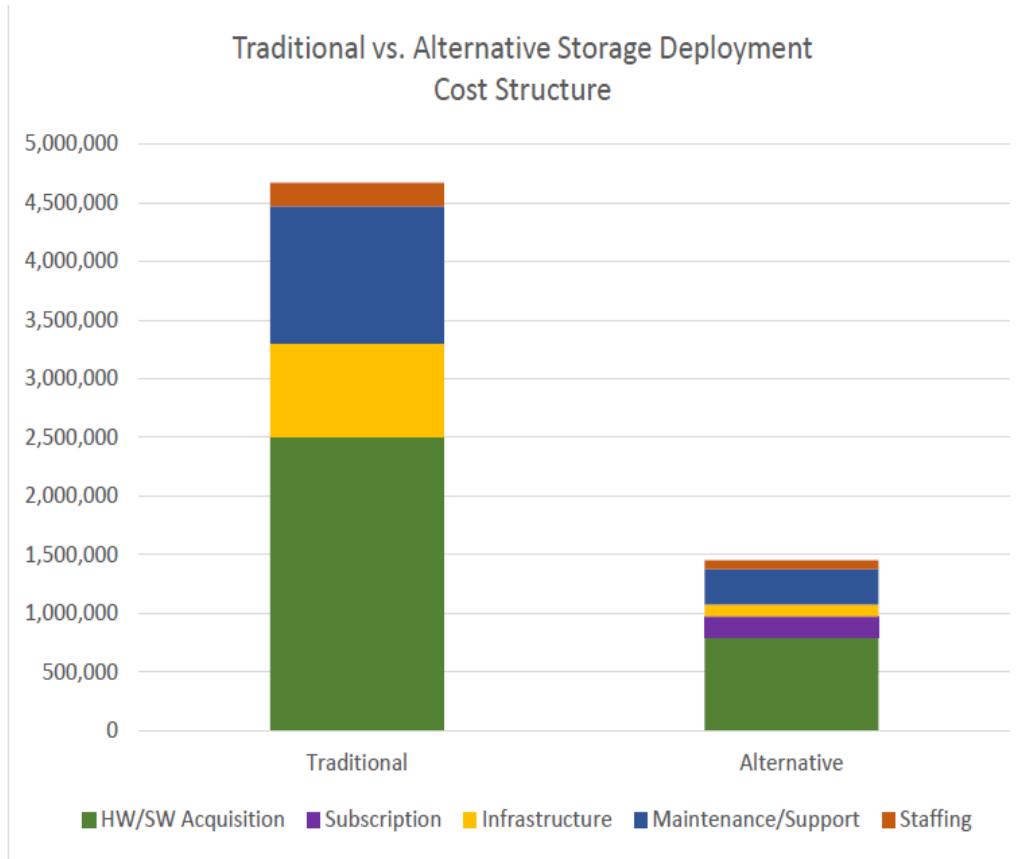
## Dla tradycyjnego centrum danych

- NFS – zarówno dla plików, jak i obiektów S3
- NFS dla S3 dostarcza mechanizm do masowego ładowania danych w istniejące repozytoria S3
- SMB/CIFS – dostęp dla klientów działających pod systemami Microsoft
- iSCSI zabezpiecza pomost do przyszłego rozwoju

## Natywnie wspierane protokoły:

- RBD (blokowy)
- RADOS (obiektowy)
- CephFS (plikowy)
- S3 oraz Swift

# Dlaczego rozwiązanie SDS jest tańsze?



"Source: Defining the Software Defined Storage Market" Marc Stalmer, Dragon Slayer Consulting

# Nacisk na łatwość zarządzania

SUSE Enterprise Storage

Dashboard OSDs RBDs Pools Nodes iSCSI NFS Object Gateway CRUSH Map System

Ceph Pools

Users Buckets

Name	ID	Used	Applications	Placement groups	Replica size	Erasure code profile	Type	Crush ruleset	Compression mode
.rgw.root	1	0.00%	rw	8	3		replicated	0	
cephfs_data	6	0.00%	cephfs	128	3		replicated	0	
cephfs_metadata	7	0.00%	cephfs	128	3		replicated	0	
default.rgw.buckets.index	28	0.00%	rw	8	3		replicated	0	
default.rgw.control	2	0.00%	rw	8	3		replicated	0	
default.rgw.log	4	0.00%	rw	8	3		replicated	0	
default.rgw.meta	3	0.00%	rw	8	3		replicated	0	
iSCSI-Images	5	0.00%	rd	128	3		replicated	0	
vmwarecert	9	0.08%	rd	4096	3		replicated	0	
vmwaretest	8	0.88%	rd	2048	3		replicated	0	

Showing 1 to 10 of 10 items

SUSE Enterprise Storage

Dashboard OSDs RBDs Pools Nodes iSCSI NFS Object Gateway CRUSH Map System

Last 12 hours Refresh every 30s

Health Status: OK

Monitors In Quorum: 3

Pools: 13

Cluster Capacity: 796 TiB

Used Capacity: 15.36 TiB

Available Capacity: 98.1%

OSDs IN: 147

OSDs OUT: 0

OSDs UP: 147

Average PGs per OSD: 134

Average OSD Apply Latency: 104.7 µs

Average OSD Commit Latency: 104.7 µs

Average Monitor Latency: 0 s

Capacity: 796 TiB

IOPS: 22:00 00:00 02:00 04:00 06:00 08:00

Throughput: 22:00 00:00 02:00 04:00 06:00 08:00

Legend: Available (Yellow), Used (Red), Total Capacity (Blue)

Legend: Write (Green), Read (Orange)

# Demo!

# SUSE Enterprise Storage – Mapa drogowa

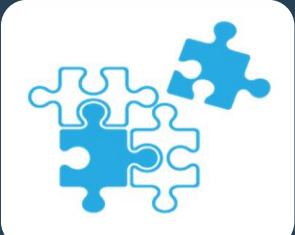
SES 4: Nov 2016

SES 5: Oct 2017

SES 5.5: Oct 2018

SES 6: June 2019\*

SES 7: Q2/Q3 2020\*



## Unified File, Object & Block

- CephFS
- Multisite object replication
- Acquired openATTIC

## Performance and Ease of Use

- BlueStore
- Compression
- Block/File Erasure code
- NFS/S3 interface
- NFS/CephFS interface
- OpenATTIC integrated with DeepSea
- Enhanced monitoring and management

## Interoperability and Security

- Integration with SUSE OpenStack Cloud
- CIFS/Samba interface
- Non SUSE Ceph clients
- Event notification
- AppArmor integration
- Security audit testing

## Enhanced Management and Hybrid Cloud

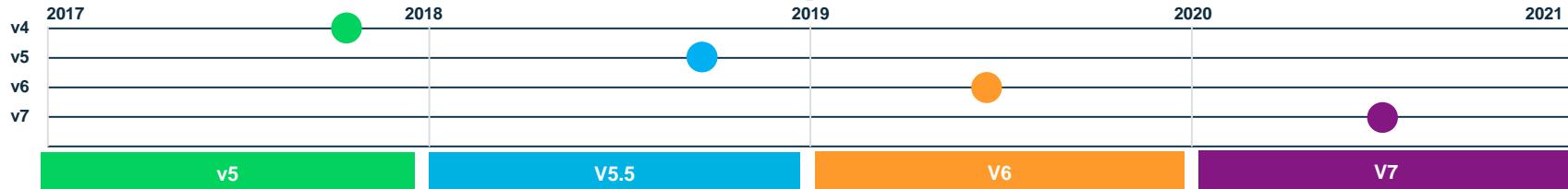
- Ceph Dashboard
- Enhanced monitoring
- Public cloud product
- Sync to external cloud
- Phone home
- RBD QoS
- CephFS snapshot
- Async file replication

## Containerization and Autonomous Self-Healing

- Containerization
- Enhanced intelligent data placement
- Guided troubleshooting
- Guided HW upgrade
- Guided OSD repair
- Data deduplication
- RDMA interconnect
- Windows drivers

\*: These dates are indicative only, not guaranteed

# SUSE Enterprise Storage



## v5

Built On
<ul style="list-style-type: none"> <li>Ceph Luminous release</li> <li>SUSE Linux Enterprise Server 12 SP3</li> </ul>
Manageability
<ul style="list-style-type: none"> <li>openATTIC phase 2           <ul style="list-style-type: none"> <li>Grafana monitoring dashboard</li> <li>Prometheus event alert - email</li> </ul> </li> <li>DeepSea (Salt) phase 2           <ul style="list-style-type: none"> <li>Online Filestore to BlueStore</li> </ul> </li> </ul>
Interoperability
<ul style="list-style-type: none"> <li>NFS Ganesh</li> <li>NFS access to S3 buckets</li> <li>CIFS Samba**</li> <li>CephFS Multi MDS support</li> </ul>
Availability
<ul style="list-style-type: none"> <li>Erasure coded block and file</li> </ul>
Efficiency
<ul style="list-style-type: none"> <li>BlueStore back-end</li> <li>Data compression</li> </ul>

## v5.5

Built On
<ul style="list-style-type: none"> <li>Ceph Luminous release</li> <li>SUSE Linux Enterprise Server 12 SP3</li> </ul>
Manageability
<ul style="list-style-type: none"> <li>Internationalization</li> <li>Usability enhancements</li> <li>Predefined profiles</li> <li>Autonomous data balancer**</li> <li>DeepSea (Salt) phase 3</li> </ul>
Interoperability
<ul style="list-style-type: none"> <li>Non SUSE RBD and CephFS clients</li> <li>CIFS/Samba</li> <li>AppArmor security module</li> </ul>
Availability
<ul style="list-style-type: none"> <li>Multisite RADOSGW N+1 with N/N-1</li> </ul>
Efficiency
<ul style="list-style-type: none"> <li>BlueStore/RocksDB optimizations</li> </ul>

## v6

Built On
<ul style="list-style-type: none"> <li>Ceph Nautilus release</li> <li>SUSE Linux Enterprise Server 15 SP1</li> </ul>
Manageability
<ul style="list-style-type: none"> <li>Ceph Dashboard (oA replacement)</li> <li>Phone home (metrics and error analysis)</li> <li>Enhanced PG balancing</li> <li>Autonomous PG scaling**</li> <li>CephFS directory quotas</li> <li>Graceful system shutdown</li> <li>Storage enclosure disk integration</li> </ul>
Interoperability
<ul style="list-style-type: none"> <li>IPv6</li> <li>RGW Elasticsearch sync module</li> <li>Node to node encryption**</li> <li>Multiple CephFS instances in one cluster**</li> <li>NFS-Ganesha active/active architecture**</li> </ul>
Availability
<ul style="list-style-type: none"> <li>Unidirectional sync external cloud via S3</li> <li>CephFS snapshots</li> <li>Asynchronous file replication**</li> </ul>
Efficiency
<ul style="list-style-type: none"> <li>QoS for RBD and background operations</li> <li>S3 object policy driven tiering**</li> </ul>

## v7

Built On
<ul style="list-style-type: none"> <li>Ceph Octopus release</li> <li>SUSE CaaS Platform</li> </ul>
Manageability
<ul style="list-style-type: none"> <li>Integration with Rook + Kubernetes</li> <li>Phone home phase 2</li> <li>Autonomous PG scaling</li> <li>Self-healing PG</li> </ul>
Interoperability
<ul style="list-style-type: none"> <li>Containerized deployment</li> <li>Native Windows client drivers</li> <li>Node to node encryption</li> <li>Multiple CephFS instances in one cluster</li> <li>NFS-Ganesha active/active architecture</li> <li>RDMA back-end**</li> </ul>
Availability
<ul style="list-style-type: none"> <li>Bidirectional sync external cloud via S3</li> <li>Asynchronous file replication</li> </ul>
Efficiency
<ul style="list-style-type: none"> <li>Data deduplication</li> <li>SW cache layer</li> <li>Object Storage Daemon optimizations</li> <li>S3 object policy driven tiering</li> </ul>

\*\* Items are tech preview

\* Information is forward looking and subject to change at any time.

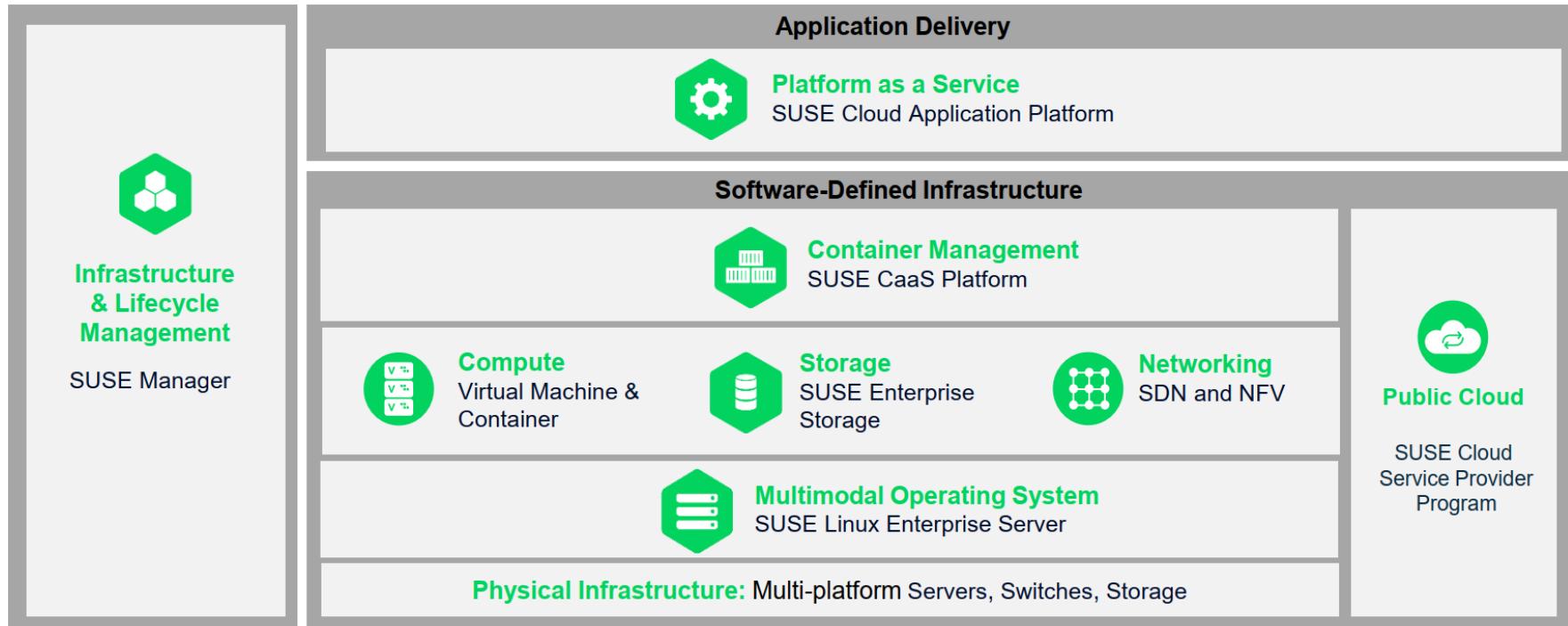
# SUSE Enterprise Storage — Zarządzanie

v4	v5	v5.5	v6	v7
2017	2018	2019	2020	2021
v5	v5	v5.5	v6	v7
<b>Install</b> <ul style="list-style-type: none"><li>Import unmanaged SUSE Enterprise Storage cluster</li></ul> <b>Configure</b> <ul style="list-style-type: none"><li>Online Filestore to BlueStore</li></ul> <b>Manage</b> <ul style="list-style-type: none"><li>Deploy, configure, assign nodes</li><li>iSCSI target management</li><li>RGW users, buckets, keys</li><li>View, create, delete CephFS</li><li>NFS Ganesh gateway</li></ul> <b>Monitor</b> <ul style="list-style-type: none"><li>Grafana dashboards</li><li>CephFS status</li></ul> <b>Diagnose</b> <ul style="list-style-type: none"><li>OSD health status</li></ul>	<ul style="list-style-type: none"><li><b>Configure</b></li><li>Convert EC to replication (visa versa)</li></ul> <ul style="list-style-type: none"><li><b>Manage</b></li><li>RBD snapshot</li><li>OpenStack integration</li><li>Cluster rebuild progress</li><li>Admin node redundancy phase 1</li></ul> <ul style="list-style-type: none"><li><b>Monitor</b></li><li>Event notification alerts**</li></ul>	<b>Configure</b> <ul style="list-style-type: none"><li>iSCSI</li><li>NFS-Ganesha</li><li>Phone home</li><li>RBD QoS parameters</li><li>Background operation parameter (QoS)</li><li>RBD remote replication</li></ul> <b>Manage</b> <ul style="list-style-type: none"><li>Role-based Ceph Dashboard</li><li>Localized Ceph Dashboard</li><li>Ceph Dashboard SSO (SAMLv2)</li><li>Crush map viewer</li><li>RBD snapshot</li></ul> <b>Monitor</b> <ul style="list-style-type: none"><li>Real-time performance metrics</li><li>Scale-out Prometheus monitoring</li><li>Event notification alerts</li><li>SNMP traps</li><li>Command audit log</li></ul>	<b>Configure</b> <ul style="list-style-type: none"><li>Deploy new OSD to existing node</li><li>CIFS/Samba</li><li>RGW and CephFS remote replication</li><li>Sync to external cloud</li></ul> <b>Manage</b> <ul style="list-style-type: none"><li>CephFS snapshot</li><li>Cluster topology map</li><li>Guided upgrade process</li></ul> <b>Monitor</b> <ul style="list-style-type: none"><li>OSD and MDS statistics</li><li>Ceph client performance</li></ul> <b>Diagnose</b> <ul style="list-style-type: none"><li>Predictive disk failure analysis</li><li>Guided troubleshooting</li><li>Guided OSD repair process</li></ul>	

\*\* Items are tech preview

\* Information is forward looking and subject to change at any time.

# Przestrzeń dla cyfrowej transformacji



*Otwarte, bezpieczne, sprawdzone*





## **Unpublished Work of SUSE LLC. All Rights Reserved.**

This work is an unpublished work and contains confidential, proprietary and trade secret information of SUSE LLC. Access to this work is restricted to SUSE employees who have a need to know to perform tasks within the scope of their assignments. No part of this work may be practiced, performed, copied, distributed, revised, modified, translated, abridged, condensed, expanded, collected, or adapted without the prior written consent of SUSE. Any use or exploitation of this work without authorization could subject the perpetrator to criminal and civil liability.

## **General Disclaimer**

This document is not to be construed as a promise by any participating company to develop, deliver, or market a product. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. SUSE makes no representations or warranties with respect to the contents of this document, and specifically disclaims any express or implied warranties of merchantability or fitness for any particular purpose. The development, release, and timing of features or functionality described for SUSE products remains at the sole discretion of SUSE. Further, SUSE reserves the right to revise this document and to make changes to its content, at any time, without obligation to notify any person or entity of such revisions or changes. All SUSE marks referenced in this presentation are trademarks or registered trademarks of SUSE LLC. in the United States and other countries. All third-party trademarks are the property of their respective owners.