

Zygmunt BOK

Szpital Specjalistyczny im Prof. E. Michałowskiego, MEDHOLDING S.A.

## HETEROGENICZNE DYSKI WIRTUALNE W ŚRODOWISKU VMWARE ESXI

**Streszczenie.** W niniejszym artykule przedstawiono koncepcję heterogenicznych dysków wirtualnych, rozpostartych w zbiorze pamięci masowych typu *SAN* i *NAS*, współpracujących z maszynami wirtualnymi w środowisku *VMWARE ESXi*, na przykładzie fragmentu specjalistycznego szpitalnego Systemu Radiografii Pośredniej *Synapse* f-my FUJIFILM Medical Systems USA, Inc., używanego w Szpitalu Specjalistycznym im. Prof. E. Michałowskiego MEDHOLDING S.A., służącego do gromadzenia oraz prezentacji rentgenogramów (RTG) i tomogramów (TK). Na gruncie tej koncepcji wykazano, że możliwe jest rozszerzanie dysku wirtualnego systemu *Synapse* do postaci heterogenicznego dysku wirtualnego i zapewnienie prawidłowego funkcjonowania tego systemu, w sytuacji stałego dopływu nowych rentgenogramów i tomogramów.

## HETEROGENEOUS VIRTUAL DISKS IN VMWARE ESXI ENVIRONMENTS

**Summary.** This article presents the concept of heterogeneous virtual disks, outstretched hands in the mass storage *SAN* and *NAS* type collection, work with virtual machines in *VMWARE ESXi*, on the example of the portion of the specialized hospital System of indirect Radiography *Synapse* FUJIFILM Medical Systems USA, Inc., which is used in Specialized Hospital named Prof. E. Michalowski MEDHOLDING S.A., for the collection and presentation of the rentgenograms (x-ray) and tomographs (TK). On the basis of this concept, it has been shown that it is possible to extend the virtual disk from *Synapse* system to form a heterogeneous virtual disk and to ensure the proper functioning of the system in a situation a steady supply of new rentgenograms (x-ray) and tomographs.

## 1. Wprowadzenie

Eksploatowany w Szpitalu Specjalistycznym im. Prof. E. Michałowskiego MEDHOLDING S.A. zastany System Radiografii Pośredniej *Synapse* klasy *enterprise* f-my *FUJIFILM Medical Systems USA, Inc.* [1,2], współpracujący z bazą danych Oracle i służący do gromadzenia, dystrybucji oraz prezentacji zdjęć RTG i Tomografii Komputerowej, zainstalowany został w postaci maszyny wirtualnej działającej na platformie *VMWARE ESXi* [3,4], posadowionej na serwerze f-my *DELL*, z dwoma macierzami *RAID* [5,6]. Zainstalowano go w obszarach *Datastore1* oraz *Datastore2*, co przedstawiono na Rys. 1.

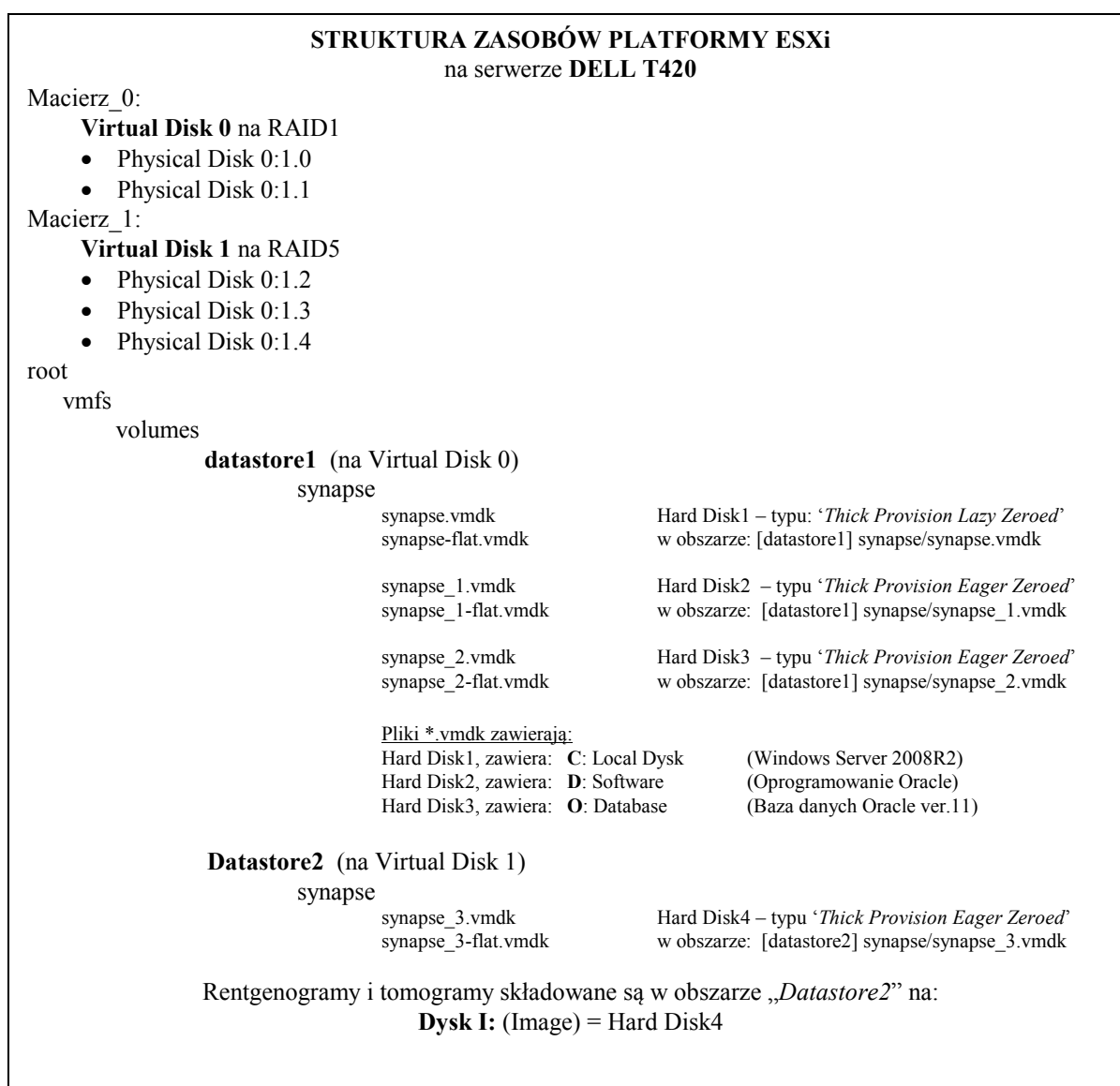


Fig. 1. *Synapse* - internal structure of resources

Dla zapewnienia prawidłowego funkcjonowania tego systemu, w sytuacji stałego dopływu nowych danych w postaci rentgenogramów i tomogramów, wymaga on ciągłego po-

większania pojemności dysku wirtualnego na którym są one składowane. Uwzględniając konieczność spełnienia ustawowego obowiązku zachowywania ich przez okres 20 lat (dla konkretnego pacjenta od ostatniego jego badania) powiększanie jego pojemności nie jest zadaniem trywialnym, w szczególności w kontekście nieuchronnych długotrwałych przestoju w pracy systemu, uniemożliwiających normalną diagnostykę medyczną. Jedną z metod, pozwalających zwiększyć pojemność tego wirtualnego dysku, jest powiększenie przestrzeni dyskowej w macierzy *RAID5*, poprzez wymianę starych dysków na dyski nowe o większej pojemności i powtórna instalacja całego systemu, kończąca się długotrwałym procesem importowania do tego dysku wirtualnego rentgenogramów i tomogramów z wcześniej wykonanej kopii. Kolejną metodą może być migracja tego systemu do pamięci masowej innych producentów wspieranych przez firmę *VMware Inc.* [7], pociągająca jednak za sobą zapewnienie nakładów inwestycyjnych, uruchomienie czasochłonnego procesu inwestycyjnego oraz nieuchronny długotrwały postój tego systemu. Wśród innych metod [8,9] pozwalających zwiększyć pojemność dysków wirtualnych można wskazać na metodę [10], która pozwala zwiększyć pojemność wirtualnego dysku w sytuacji, kiedy wielkość partycji w fizycznym systemie serwerowym nadal pozostaje taka sama. W celu zmiany wielkości partycji muszą zostać przeprowadzone specyficzne kroki dla danego serwera i systemu operacyjnego. Dla serwerów firmy *Dell* istnieje metoda [11,12], wykorzystująca oprogramowanie *Dell OpenManage Server Administrator vSphere Installation Bundle (VIB) for ESXi*, pozwalająca na zwiększenie pojemności tego wirtualnego dysku, której istota polega na rozszerzeniu przestrzeni dyskowej aktualnej macierzy *RAID5* bez utraty danych. Dla serwerów firmy *IBM* dostępne jest oprogramowanie *EaseUS Partition Master Server Edition*, pozwalające na rozszerzenie, z poziomu *Windows Server*, partycji *RAID5*. Po uzyskaniu dodatkowej przestrzeni dyskowej, w postaci powiększonego istniejącego woluminu zdefiniowanego w aktualnej macierzy *RAID5*, można zwiększyć pojemność dysku wirtualnego o kolejne jednostki alokacyjne, tj.: do 2 TB - w przypadku *VMWARE ESXi* ver. 5.5 oraz powyżej 2 TB - w przypadku *VMWARE ESXi* ver. 6.0, z zasobów powiększonego istniejącego woluminu.

W niniejszym artykule przedstawiono metodę pozwalającą zwiększyć pojemność dysku wirtualnego, przy jednoczesnym radykalnym skróceniu czasu przestoju systemu informatycznego spowodowanego tą operacją, w porównaniu z czasami przestoju spotykanymi w przypadku innych metod. Przedstawiona metoda opiera się na koncepcji heterogenicznych woluminów dysków wirtualnych, przedstawionych na Rys. 2, rozpostartych na zbiorze heterogenicznych pamięci masowych typu *SAN* i *NAS*, współpracujących z maszynami wirtualnymi w środowisku *VMWARE ESXi*. Na gruncie tej koncepcji wykazano, że możliwe jest rozszerzanie dysku wirtualnego systemu *Synapse* do postaci

heterogenicznego dysku wirtualnego i zapewnienie prawidłowego funkcjonowania tego systemu, w sytuacji stałego dopływu nowych rentgenogramów i tomogramów. Rozszerzenie dysku wirtualnego z maszyny wirtualnej systemu *Synapse* przechowujących rentgenogramy i tomogramy dokonano za pomocą oprogramowania *VMware vSphere* [13]. W istniejącej maszynie wirtualnej, utworzono nowy dysk wirtualny w nowym obszarze typu *Datastore* w kolejnej pamięci masowej, a następnie, z poziomu systemu operacyjnego tej wirtualnej maszyny, istniejący dysk wirtualny rozpostarto na nowo utworzony dysk. Rozszerzenia dysku dokonywano za pomocą narzędzia administracyjnego „*Computer Management*”, tj. *Disk Management* z opcją *Extend Volume*.

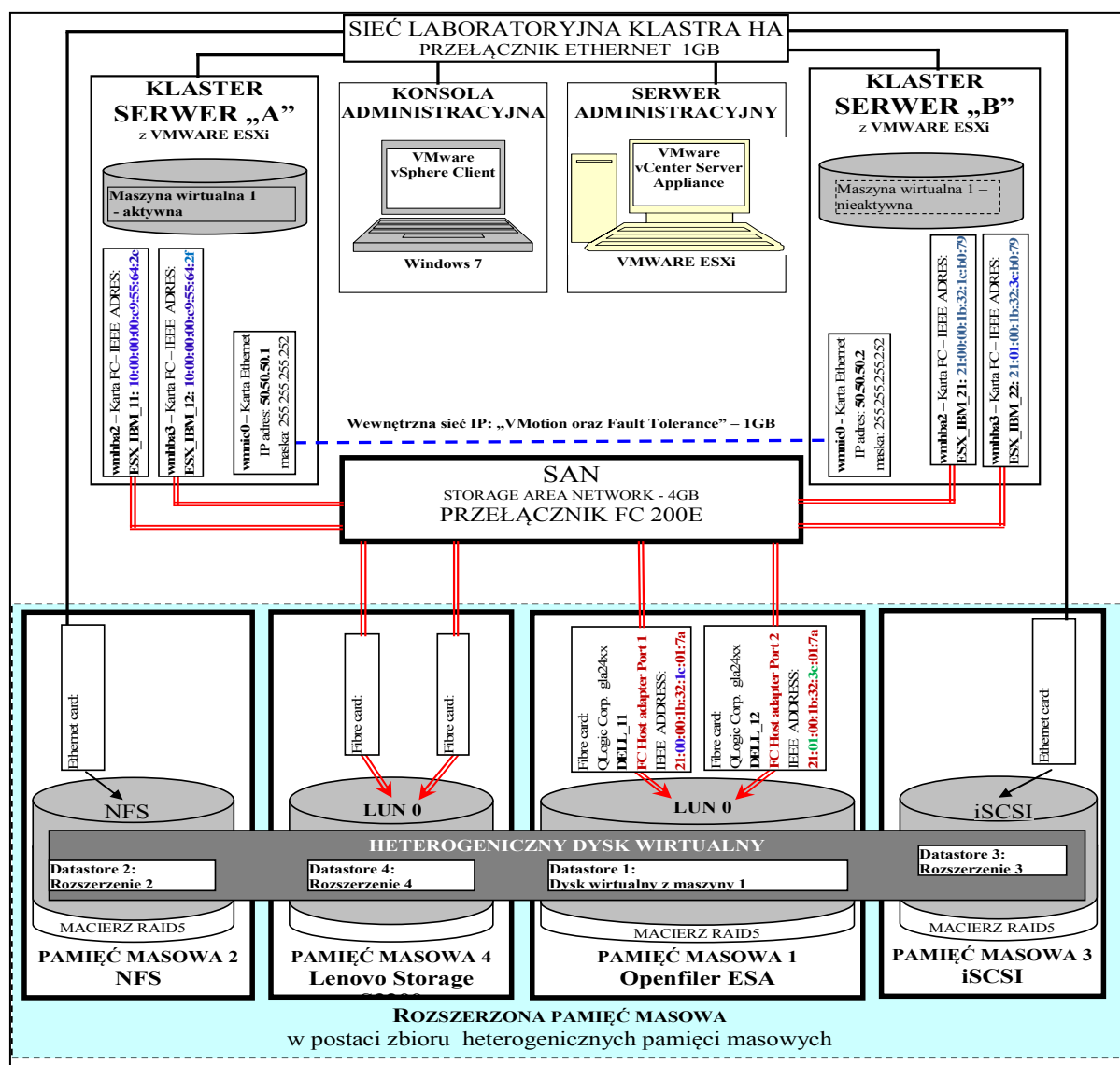


Fig. 2. Heterogeneous virtual disk spread on different type mass storage

Celem pracy było zbadanie w środowisku badawczym, w postaci laboratoryjnej infrastruktury klastrowej przedstawionej na Rys. 2., możliwości:

1. uruchomienia fragmentu specjalistycznego szpitalnego Systemu Radiografii Pośredniej *Synapse* w pamięci masowej typu *block-level Fiber Channel SAN Storage* [13,14,15,16,17,18,19,20,21,22,23], działającą pod kontrolą systemu operacyjnego *Linux Openfiler ESA* [24,25,26,27,28,29,30] zapewniającego obsługę plikowych (*file-level*) typu *NAS (Network-Attached Storages)* oraz blokowych (*blok-level*) pamięci masowych typu *SAN Storage*, współpracującą z klastrem wysokiej dostępności *HA (High Availability) VMWARE ESXi* w zakresie podstawowej zdolności przyłączeniowej *BC (Basic Connectivity)*, odporności na uszkodzenia *FT (Fault Tolerance)* oraz pełnej redundancji połączeń światłowodowych pomiędzy pamięcią masową a klastrem *HA* zbudowanym w technologii *VMWARE ESXi*,
2. rozszerzania aktualnego lokalnego dysku wirtualnego systemu *Synapse* z bazą danych *Oracle*, przechowującego rentgenogramy i tomogramy, do postaci heterogenicznego dysku wirtualnego rozpostartego w zbiorze heterogenicznych pamięci masowych typu *block-level Fiber Channel SAN Storage* działającej pod kontrolą systemu operacyjnego *Linux Openfiler ESA*, *Lenovo Storage S3200*, *NFS* [31,32,33,34], *iSCSI* [35,36,37,38] oraz dokonania porównania szybkości działania tych dysków wirtualnych w stosunku do dysku wirtualnego znajdującego się w pamięci masowej typu *block-level* na dyskach lokalnych jednego z serwerów klastra *ESXi*.

W pracy dokonano również oceny porównawczej [39,40] szybkości działania systemu *Synapse* w sytuacji, kiedy jego systemowe dyski wirtualne z bazą danych *Oracle* zawierającą rentgenogramy i tomogramy znajdowały się:

1. w pamięci masowej typu *block-level*, alokowanej na dyskach lokalnych jednego z serwerów klastra *ESXi* i udostępnionej za pomocą protokołu *SCSI*,
2. w pamięciach masowych *SAN*, udostępniane za pomocą protokołu *FC (Fibre Channel)*, alokowane w pamięci masowej typu *block-level*:
  - *Fiber Channel SAN Storage* działającej pod kontrolą *Openfiler ESA*,
  - *Lenovo Storage S3200*,
3. w pamięciach masowych *NAS*, włączonych do laboratoryjnej sieci *Ethernet*, alokowane w pamięci masowej typu:
  - *block-level iSCSI* i udostępniane za pomocą protokołu *FCoE (Fibre Channel over Ethernet)*,
  - *file-level* z systemem plików *NFS* i udostępniane za pomocą protokołu *NFS (Network File System)*.

## 1.1. Specyfika przestrzeni dyskowych w systemach pamięci masowych

Pamięć masowa (*mass storage*) określa różne techniki i urządzenia pozwalające na trwałe przechowywanie dużych ilości danych cyfrowych. Istnieją dwie fundamentalne abstrakcje pojęciowe dotyczące plików (*files*) i bloków (*blocks*), umożliwiające dostęp do danych zawartych w przestrzeniach dyskowych pamięci masowych. Specyfika przestrzeni dyskowych w systemach pamięci masowych związana jest z zaimplementowaną technologią pamięci masowych, organizacją dostępu do danych typu: *BBDA* (*Block-Based Data Access*), *FBDA* (*File-Based Data Access*), *OBDA* (*Object-Based Data Access*), *ISN* (*Intelligent Storage Nodes*) [41], jak również protokołami dostępowymi typu *File-Level* oraz *Block-Level*, udostępniających przestrzeń dyskową w systemach pamięci masowych [42,43] typu:

- *File-Level*, rozpowszechniona w pamięciach masowych typu *NAS* [44] (*Network Attached Storage*) – technologia umożliwiająca podłączenie zasobów pamięci dyskowych bezpośrednio do sieci komputerowej,
- *Block-Level*, rozpowszechniona w pamięciach masowych typu *DAS* (*Direct Attached Storage*) [45,46] oraz *SAN* (*Storage Area Network*) [47,48] - technologia zapewniająca systemom komputerowym dostęp do zasobów pamięci masowej.

### 1.1.1. Specyfika przestrzeni dyskowych o organizacji plikowej

Pamięć masowa typu *File-Level* jest popularnym systemem pamięci masowej, którą można odnaleźć w dyskach twardych komputerów oraz dyskach *NAS*, które są serwerowymi magazynami danych połączonych z siecią *Ethernet*, zapewniającymi współdzielony dostęp plików heterogenicznym klientom [49]. Urządzenie pamięci masowej obsługuje pliki i foldery znajdujące się w niej. Oznacza to, że w wielu przypadkach, urządzenie pamięci masowej *File-Level* lub *NAS* musi obsługiwać kontrolę ustawień dostępu użytkowników i ich uprawnień. Niektóre z urządzeń integrują się z istniejącym systemem autentykacji i bezpieczeństwa. Są proste w implementacji i użyciu; magazynują pliki i foldery, które są widoczne i dostępne dla systemów operacyjnych. Ponadto obsługują kontrolę dostępu oraz integrację z korporacyjnymi usługami katalogowymi.

### 1.1.2. Specyfika przestrzeni dyskowych o organizacji blokowej

Systemy pamięci masowej typu *Block-Level*, biorąc pod uwagę elastyczność i wszechstronność, mają zdecydowaną przewagę nad systemami pamięci masowych typu *File-Level*. W urządzeniach masowych typu *Block-Level* tworzone są surowe bloki lub woluminy magazynowe, a następnie serwerowy system operacyjny łączy się z tymi blokami, traktując je

i kontrolując jako indywidualne dyski twarde. Każdy blok lub wolumin magazynowy może być indywidualnie formatowany wymaganym systemem plików np. *NFS (Linux, VMware)*, *NTFS*, *SMB/CIFS (Windows)*, *VMFS (Virtual Machine File System – VMware)*, itp. To czyni te pamięci użyteczne dla prawie każdego rodzaju aplikacji, w tym przechowywanie plików i baz danych. Główną zaletą stosowania systemów pamięci masowej typu *Block-Level* jest lepsza wydajność/prędkość w stosunku do pamięci masowych typu *File-Level*. Transport danych jest bardziej efektywny i niezawodny. Zabezpieczenie danych w postaci *backup*'u [50] realizowane jest za pomocą natywnych, dla danego systemu operacyjnego, narzędzi *backup*'owych, chociaż istnieją również narzędzia firm trzecich.

## 1.2. Protokoły udostępniające przestrzeń dyskową

Dostęp do przestrzeni dyskowych w systemach pamięci masowych w nierozdzielny sposób związany jest z protokołami [51] typu *File-Level* oraz *Block-Level*, udostępniających przestrzeń dyskową w tych systemach.

### 1.2.1. Protokoły udostępniające przestrzeń dyskową o organizacji plikowej

Pamięć masowa typu *File-Level* może być konfigurowana i udostępniana za pomocą protokołów typu *File-Sharing*, takich jak: *NTFS*, *SMB/CIFS (Windows)*, *NFS (Unix, Linux, VMware)*, *NCP (NetWare)*, w celu zapewnienia widoczności i dostępności dla systemu operacyjnego. Protokoły typu *File-Sharing* są sieciowymi protokołami wysokiego poziomu, zapewniające język i składnię dla żądań plikowych (*file requests*), pomiędzy klientami i serwerami, dostarczające poleceń dla otwierania, czytania, zapisu oraz zamykania plików poprzez sieć komputerową. Zwane są też protokołami *client/server*. Funkcjonują w warstwie 7 modelu *OSI*. W związku z zapewnieniem klientom dostępu do wielu serwerów pracującym pod kontrolą różnych systemów operacyjnych, albo klient wspiera protokół typu *File-Sharing* dla każdego systemu operacyjnego, albo serwer wspiera protokół typu *File-Sharing* dla każdego klienta. Oprogramowanie, które zawiera takie możliwości jest bardzo popularne i wspiera współdziałanie (*interoperability*) pomiędzy platformami *Windows*, *Macintosh*, *NetWare*, *Unix*, *Linux*.

*NTFS - (New Technology File System)* – to standardowy system plików systemu *Windows NT* i jego następców, wywodzący się od systemu plików *HPFS (Microsoft)* i *OS/2 (IBM)*, wprowadzony w celu zastąpienia starszego *FAT*-u oraz zapewnienia niezawodności, dodatkowych funkcji, jak również obsługę wymagań *POSIX*.

*SMB* - (*Server Message Block*) jest protokołem typu klient-serwer, oparty na systemie zapytań generowanych przez klienta i odpowiedzi od serwera. Wykorzystuje do działania dwa protokoły niższych rzędów - protokół warstwy sesji *NetBIOS* lub protokół *NetBEUI*.

*CIFS* - (*Common Internet File System*) - jest protokołem transferowym typu *Shared Storage*, używanym w środowisku *Microsoft* dla współdzielenia plików [52], wspierający system magazynowania danych typu *File-Level*, oparty na protokole *SMB*, zapewniający ujednoczony dostęp do plików, niezależny od systemu operacyjnego, u podstaw którego kryją się inne systemy plików takie jak: *FAT32*, *XFS*, *NTFS*.

*NFS* - (*Network File System*) [53] - protokół typu *File-Sharing*, opracowany przez *Sun Microsystems*, pełni rolę standardu dla współdzielenia plików wśród hostów *UNIX*owych, służący do zdalnego udostępniania systemu plików, łączący systemy heterogeniczne z klientami i serwerami pracującymi pod kontrolą różnych systemów operacyjnych, posadowionych na różnym sprzęcie.

*NCP* - (*NetWare Core Protocol*) – protokół sieciowy, opracowany i wykorzystywany przez firmę *Novell*, używany pod kontrolą systemu operacyjnego *NetWare*, jako główny protokół warstwy 7 sieci *NetWare*, zapewniający usługi związane z plikami, drukowaniem oraz kolejkowaniem.

### **1.2.2. Protokoły udostępniające przestrzeń dyskową o organizacji blokowej**

Pamięci masowe typu *Block-Level* używają protokołów *SCSI* (*Small Computer Systems Interface*), *iSCSI* (*Internet SCSI*), *FC* (*Fibre Channel*), *FCoE* (*Fibre Channel over Ethernet*) dla transferu danych w sieciach komputerowych jak polecenia *SCSI*, działające jako interfejs komunikacyjny pomiędzy „inicjatorem” i „targetem”. Ze względów historycznych i porządkowych, należy wspomnieć też o protokołach *ATA* (*Advanced Technology Attachments*), które zasadniczo nie są używane w sieciach pamięci masowych, ale dyski *ATA* są wykorzystywane w podsystemach pamięci masowych, które łączą się z tymi sieciami. W przestrzeniach dyskowych o organizacji blokowej, pamięć masowa typu *Block-Level* prezentuje siebie serwerom używającym mechanizm połączeniowy w postaci przemysłowego standardu *FC* oraz *iSCSI*.

#### **1.2.2.1. Protokół ATA**

*ATA* - jest protokołem pamięci masowych typu *open-systems*, ustandaryzowany przez Komitet *ANSI X3*. Kilka początkowych standardów *ATA*, opublikowanych przez *ANSI ATA*, specyfikują protokoły typu *block-level* wspierające tylko polecenia dyskowe. Wraz z wpro-



wadzeniem standardu *ATAPI*, wprowadzono polecenia podobne do protokołu *SCSI* (*SCSI-like commands*).

#### 1.2.2.2. Protokół SCSI

*SCSI* – protokół typu *block-level*, jest standardem protokołu polecenia transportu bloku (*block transport command protocol*), używany dla zapewnienia wysokiej przepustowości transportu danych pomiędzy hostami i systemami pamięci masowej [54,55,56].

#### 1.2.2.3. Protokół iSCSI

*iSCSI* – to technika umożliwiająca wykonywanie operacji wejścia-wyjścia na dysku twardym odległej maszyny za pomocą protokołu *TCP/IP*. Protokół *iSCSI* [57,58,59] umożliwia budowę systemów pamięci masowych *SAN*, przy zastosowaniu macierzy dyskowych *SCSI* i sieci *Ethernet*. Protokół *iSCSI*, obsługiwany przez systemy *Windows* i *Linux*, jest protokołem typu *Block-Level*, który hermetyzuje (enkapsuluje) polecenia *SCSI* do pakietów *TCP/IP*.

#### 1.2.2.4. Protokół FC

*FC* – jest ogólną nazwą zintegrowanego zbioru standardów opracowanych przez *ANSI* [60,61,62,63]. Protokół *FC*, mogący obsłużyć każdą topologię i protokoły transportowe *IP* oraz *SCSI*, został zaprojektowany przez *ANSI* w standardzie *T11* Komitetu Technicznego *INCITS*, w celu zapewnienia łączności, niezawodności oraz znaczącego poprawienia prędkości transferu danych pomiędzy stacjami roboczymi, komputerami i urządzeniami magazynującymi dane.

#### 1.2.2.5. Protokół FCoE

*FCoE* – jest protokołem hermetyzującym ramki protokołu *FC* do ramek *Ethernet*. Specyfikacja *FCoE* jest częścią standardu *T11 FC-BB-5*, opracowanego przez *INCITS* [64,65].

### 1.3. Kompresja przestrzeni dyskowej

#### 1.3.1. Kompresja danych w przestrzeni dyskowej na poziomie organizacji plikowej

W przeciwieństwie do kompresji plików typu *file compression*, zapewnianych przez formaty archiwizacyjne takie jak: *zip*, *bzip2*, *gzip*, w których kompresja plików nie może być opisana jako transparentna, wiele systemów operacyjnych oferuje kompresję za pomocą oprogramowania typu *FDU* (*Fly-Disk-Utility*) lub przezroczystą kompresję na poziomie *file-level*. Pozwala ona wszystkim programom i narzędziom informatycznym na skorzystanie z niej, bez konieczności stosowania specjalistycznego oprogramowania do kompresji plików.

Korzyści wynikające z transparentnej kompresji, to przede wszystkim oszczędność przestrzeni dyskowej oraz zredukowana ilość operacji dyskowych *I/O*, które są najwolniejszymi operacjami komputerowymi - o wiele rzędów wielkości wolniejsze niż instrukcje *CPU*. Dlatego też minimalizacja operacji *I/O* oferuje większy potencjał na poprawę ogólnej wydajności systemu komputerowego. Mniej operacji *I/O* oznacza szybsze ładowanie do pamięci operacyjnej skompresowanego pliku w porównaniu z plikiem nieskompresowanym. Oszczędność przestrzeni dyskowej oraz operacji *I/O* jest jednak obciążona kosztami. Kiedy plik jest zapisywany, to poddawany jest kompresji; przy czytaniu, poddawany jest dekompresji. To powoduje wzrost obciążenia *CPU*, w szczególności, kiedy zastosowano złożone algorytmy kompresji typu *VSE* (*Very Space Efficient*).

Obok historycznie pierwszych systemów oferujących kompresję typu *fly-disk-utility*, za pomocą oprogramowania typu *DCU* (*Disk Compression Utility*), bardzo popularne w wczesnych latach dziewięćdziesiątych z systemem *Caldera DR-DOS 7.02*, zawierającym kompresję *DoubleSpace*, do nowoczesnych systemów operacyjnych oferujących przezroczystą kompresję na poziomie *file-level* można zaliczyć system plików *HFS+* [66]. Zawiera on transparentną kompresję plików typu *per file compression*, którego przykładem może być *Snow Leopard 10.6* z systemu operacyjnego *Mac OS X*, którego zestaw funkcji z systemu kompresji jest mocno ograniczony w stosunku do innych systemów plików, jak np. *NTFS*.

System plików *NTFS* używany przez *Windows* posiada wbudowaną funkcjonalność znaną jako kompresja *NTFS*. Umożliwia ona przezroczystą kompresję plików sprawiając, że zużywają one mniej przestrzeni dyskowej na dysku twardym, przy jednoczesnym normalnym dostępie do plików. Kompresja *NTFS* jest dostępna tylko na woluminach używających systemu plików *NTFS*, która może być użyta do kompresji plików i folderów. Pliki są dekompresowane automatycznie i wyglądają w systemie *Windows* tak, jak przed kompresją. Kompresja *NTFS* nie powinna być używana do: (i) systemowych plików *Windows* i innych ważnych plików programowych, (ii) serwerów z dużym obciążeniem *CPU* skutkujące zwiększeniem czasu dostępu do plików, (iii) nowoczesnych desktopów lub laptopów z *CPU* mogącym pozostawać w stanie bezczynności.

### 1.3.2. Kompresja danych w przestrzeni dyskowej na poziomie organizacji blokowej

Kompresja na poziomie *BLC (Block-Level-Compress)* [67] umożliwia przechowywanie każdego normalnego bloku danych w formie skompresowanej. *BLC* jest metodą kompresji bezstratnej. Ma zastosowanie do podstawowych tablic danych, tj. do:

- *BLOB (Binary Large Object)* – typ danych, który umożliwia przechowywanie dużych ilości danych binarnych w postaci pojedynczego obiektu w bazie danych,
- *CLOB (Character Large Object)* - kolekcja danych znakowych w systemach zarządzania bazą danych. Obiekty *CLOB* najczęściej są przechowywane w wydzielonym miejscu, odwołanie do którego znajduje się w tabeli oraz podtablicach *CLOB*,
- tabel rezerwowych, podtabeli indeksów złączeniowych, podtabeli indeksu odniesienia i tabel słownikowych.

Celami kompresji na poziomie *block-level* jest zaoszczędzenie przestrzeni dyskowej oraz redukcja dyskowej *I/O* przepustowości. Kompresja *block-level* może znacząco zwiększyć wykorzystanie *CPU* w przypadku dynamicznej kompresji i dekompresji danych. Ponadto może zwiększyć zysk w przepustowości z powodu redukcji ruchu *I/O*. Transparentna kompresja [68,69] *block-level* może poprawić zarówno efektywność wykorzystania przestrzeni dyskowej oraz wydajność *online* magazynów pamięci masowej.

## 2. Instalacja i konfiguracja środowiska badawczego

W związku z nakreślonymi celami, zbudowano środowisko badawcze w postaci klastra badawczego, zgodnie z Rys.1, w zakresie podstawowej zdolności przyłączeniowej, odporności na uszkodzenia *FT* oraz redundancji połączeń światłowodowych pomiędzy klastrem a pamięcią masową typu „*Fiber Channel SAN Storage*”, pracującej pod kontrolą systemu *Linux Openfiler ESA*.

### 2.1. Konfigurowanie klastra ESXi

Klaster laboratoryjny zbudowano i skonfigurowano na dwóch maszynach *IBM System X3500*, na których zainstalowano oprogramowanie serwerowe *VMWARE ESXi ver. 5.5*. Zarządzanie klastrem sprawowano za pomocą pakietu oprogramowania *VMware vCenter Operations Management Suite 5.8*, tj. *VMware vCenter Server Appliance ver. 5.5*, które zainstalowano na oddzielnym serwerze zarządzającym z oprogramowaniem serwerowym

*VMware ESXi ver. 5.5*. Oprogramowanie administracyjne, w postaci pakietu *VMware vSphere Client ver.5.5*, zarządzające całym klastrem zainstalowano na kolejnym komputerze.

W celu instalacji oprogramowania zarządzającego *VMware vCenter Server Appliance*, korzystano z kreatora tego oprogramowania z poziomu klienckiego programu administracyjnego *vSphere Client*, instalującego plik *VMware-vCenter-Server-Appliance-5.5.0.10000-1624811\_OVF10.ovf* zawierający oprogramowanie zarządzające.

Dla uzyskania dostępu do konsoli administracyjnej *VMware vCenter Server Appliance*, na serwerze zarządzającym uruchomiono maszynę wirtualną dostępną pod nazwą *VMware vCenter Server Appliance*. Dostęp do uruchomionej konsoli maszyny wirtualnej *VMware vCenter Server Appliance* uzyskiwano za pomocą dowolnej przeglądarki internetowej.

Właściwą strukturę klastra utworzono za pomocą kreatora z pakietu *VMware vSphere Client for ESXi servers*, zainstalowanego na konsoli administracyjnej. W procesie tworzenia klastra ustawiono następujące jego opcje oraz funkcjonalności, tj.:

1. Włączono opcję wysokiej dostępności - *Turn On vSphere HA*,
2. Włączono opcję monitorowania hostów - *Enable Host Monitoring*,
3. Włączono wysoki priorytet restartu maszyn wirtualnych - *VM restart priority - High*,
4. Wyłączono opcję monitorowania maszyn wirtualnych - *VM Monitoring Status - Disabled*,
5. Wyłączono kompatybilność - *Enhanced vMotion Compatibility – Disable EVC*,
6. Lokalizacja pliku *swapfile*: *Store the swapfile in the same directory as the virtual machine*.

W celu zapewnienia funkcjonalności związanej z przenoszeniem wirtualnych maszyn pomiędzy hostami i klastrami (*VMotion*) oraz odporności na awarie *FT*, utworzono dla nich wewnętrzną sieć komunikacyjną pod nazwą *VMotion and Fault Tolerance*.

## **2.2. Konfigurowanie pamięci masowej typu *SAN Fibre Channel***

### **2.2.1. Część 1. *Linux Openfiler ESA – Konfigurowanie kanału światłowodowego***

Instalację systemu operacyjnego *Openfiler ESA* (ver. 2.99.1) przeprowadzono na maszynie *Dell PowerEdge 2900* z dwiema kartami światłowodowymi typu *qla24xx*, włączono usługi „*scst*” (# *chkconfig scst on*) oraz „*qla2x00tgt*” (# *chkconfig qla2x00tgt on*), odszukano porty *WWN* [2] (# *cat /sys/class/fc\_host/host\*/port\_name*), a następnie dla każdego interfejsu światłowodowego włączono tryb „*target*” (za pomocą komendy o składni: *scstadmin -enable\_target X-driver Y*, gdzie *X,Y* jest adresem *WWN*). Z kolei dla urządzeń, które będą używały trybu *target*, utworzono grupę bezpieczeństwa o nazwie „*esxi*”, za pomocą polecenia o składni: *scstadmin -add\_group Z -driver Y -target X*, gdzie *Z* jest nazwą grupy.

## 2.2.2. Konfigurowanie przełącznika Fibre Channel FC

### 2.2.2.1. Definiowanie i tworzenie aliasów oraz stref

W celu ograniczenia dostępu serwerów do pamięci masowej nie alokowanych dla nich, w sieciach SAN używa się tzw. mechanizmu *zoningu*, który pozwala na segmentację sieci typu *Fibre Channel* przy pomocy przełączników. Zazwyczaj strefy (*zones*) tworzone są dla każdej grupy serwerów, które uzyskują dostęp do udostępnionej grupy urządzeń pamięci masowej i jednostek logicznych *LUN* (*Logical Unit Number*) [70,71]. Strefy definiują, które hosty mogą łączyć się z określonymi pamięciami masowymi. W zastosowanym przełączniku *FC 200E*, zdefiniowano oraz utworzono:

- aliasy [72] dla hostów klastra i pamięci masowej, za pomocą komendy o składni:  
`# alicreate Alias_hosta, Adres_WWW`
- oraz strefy [73,74,75], za pomocą komendy o składni:  
`# zonecreate Nazwa_Strefy, "Alias_Portu_Hosta; Alias_Portu_Pamięci_Masowej"`

### 2.2.3. Przyporządkowanie portów w przełączniku FC do grup bezpieczeństwa

W dalszym etapie konfigurowania kanałów światłowodowych pamięci masowej pracującej pod kontrolą *Linux Openfiler ESA*, określono adresy *WWN* [72,73] wszystkich hostów *ESXi* włączonych do niej za pośrednictwem menadżera przełącznika *FC 200E*. Następnie, z poziomu systemu operacyjnego pamięci masowej przyporządkowano, za pomocą polecenia o składni `scstadmin -add_init W -driver y -target x -group Z`, porty przełącznika *FC 200E* do grup bezpieczeństwa dla kart światłowodowych zainstalowanych w serwerach klastra *ESXi* oraz w serwerze pamięci masowej *Openfiler ESA*.

## 2.2.4. Część 2. Linux Openfiler ESA – Utworzenie wolumenu logicznego

W celu utworzenia kolejnych obiektów niezbędnych do skonfigurowania pamięci masowej, w dalszej części tego rozdziału posługiwano się następującymi pojęciami:

1. Fizyczny Wolumen – przyporządkowanie przestrzeni dyskowej na fizycznym dysku dla wykorzystania w Grupie Wolumenowej,
2. Grupa Wolumenowa – zawiera fizyczne woluminy fizyczne, spośród których utworzono Logiczny Wolumen,
3. Jednostka logiczna *LUN* – jednostka logiczna, która przedstawiana jest hostom *ESXi*.

W celu utworzenia wolumenu logicznego o nazwie „*vol\_fc*”, zawartego w grupie wolumenowej „*my\_vg*”, określonej w pamięci masowej typu *Fiber Channel SAN Storage* i będącej pod kontrolą systemu *Openfiler ESA*, korzystano z jego graficznego interfejsu administracyjnego *GUI*.

#### 2.2.4.1. Konfigurowanie pamięci masowej typu FC Channel Storage

Przed konfiguracją pamięci masowej typu *FC Storage* pracującej pod kontrolą systemu *Openfiler ESA*, dokonano przygotowania wirtualnych dysków w macierzy dysków serwera *Dell PowerEdge 2900*, za pomocą jego *BIOS RAID* menadżera. Utworzono dwa wirtualne dyski:

1. *Virtual disk 0* – (*RAID 0, VOL0*), przeznaczony jako urządzenie */dev/sda* systemu operacyjnego *Openfiler ESA*,
2. *Virtual disk 1* - (*RAID 5, VOL1*), przeznaczony dla przyszłego wykorzystania przez klastr *ESXi* (dla definiowania obszaru danych *Datastore*), jako urządzenie „*/dev/sdb*”, dostępne również dla *iSCSI* lub kanału *FC*.

Za pomocą administracyjnego interfejsu graficznego *GUI* serwera *Openfiler ESA*, wprowadzono niezbędne informacje systemowe. Oprócz informacji o nazwie hosta, adresie *DNS*, adresie *IP* dla interfejsu *eth0*, wprowadzono też adresy przyłączonych hostów klastra *ESXi* oraz przełącznika *FC* z ustawioną flagą *Share*. Ponadto dokonano:

##### A. Utworzenia nowego fizycznego wolumenu

Fizyczny wolumen utworzono na podstawie wcześniej wykreowanego dysku „*Virtual disk 1*”, który następnie użyto do zdefiniowania Grupy Wolumenowej. Utworzony wolumen fizyczny posłużył do zdefiniowania przyszłego obszaru danych *Datastore*, dostępnego przez kanał *iSCSI* lub *FC*. Jako typ partycji nowego wolumenu wybrano *Physical volume*, określony jako *Primary*.

##### B. Utworzenie nowej grupy wolumenowej

Po utworzeniu wolumenu fizycznego, zdefiniowano nową grupę wolumenową w której będzie on rezydował. Po wprowadzeniu nazwy grupy wolumenowej i zaznaczeniu właściwego wolumenu fizycznego w którym będzie definiowana, utworzono grupę wolumenową o nazwie „*my\_vg*”.

##### C. Utworzenie wolumenu

Po utworzeniu grupy wolumenowej *my\_vg* wykreowano w niej wolumen „*vol\_fc*”. Określono jego rozmiar oraz typ jako *block iSCSI, FC*.

##### D. Uruchomienie usługi „*SCST Target*” oraz „*FC Target*”

Ostatnim elementem konfiguracji pamięci masowej typu *FC Storage* pracującej pod kontrolą *Openfiler ESA* było uruchomienie usług „*SCST Target*” oraz „*FC Target*”, które uaktywniano z poziomu strony administracyjnej systemu *Openfiler ESA*.

#### 2.2.4.2. Utworzenie urządzenia *scst*

Do utworzenia urządzenia *FC*, wskazującego na wcześniej utworzony wolumen logiczny „*vol\_fc*” użyto komendy o składni: *scstadmin -open\_dev V -handler T -attributes filename=U*, gdzie parametr *V* oznacza wcześniej utworzoną etykietą utworzonego woluminu „*vol\_fc*”, natomiast parametr *U* oznacza pełną ścieżką (*/dev/my\_vg/vol\_fc*) do logicznego wolumenu. Parametr *T (handler)* określono za pomocą komendy:

```
# scstadmin -list_handler
Collecting current configuration: done.
  Handler
  vdisk_fileio
```

W końcu, za pomocą komendy o składni:

```
# scstadmin -open_dev vol_fc -handler vdisk_fileio -attributes filename=/dev/my_vg/vol_fc
```

wykreowano urządzenia *FC*, wskazujące na wcześniej utworzony wolumen logiczny *vol\_fc*.

#### 2.2.4.3. Przypisanie wolumenu logicznego oraz *LUN* do grupy bezpieczeństwa

W celu przyporządkowania, dla kart sieciowych w pamięci masowej typu *FC Storage* pracującej pod kontrolą *Openfiler ESA*, wcześniej utworzonego wolumenu logicznego „*vol\_fc*” oraz *LUN* do grupy bezpieczeństwa, korzystano z komend o następującej składni: *scstadmin -add\_lun S -driver Y -target W -group Z -device V*, gdzie *S* - oznacza numer *LUN*.

#### 2.2.5. Definiowanie obszarów danych dla serwerów *ESXi* w pamięci masowej typu *SAN Fibre Channel*

Po dokonaniu konfiguracji pamięci masowej typu *SAN Fibre Channel*, zdefiniowano dla danych odpowiednie obszary typu *Datastores* dla każdego serwera *ESXi* w klastrze. Na koniec prac konfiguracyjno-instalacyjnych środowiska badawczego utworzono obszar pamięci masowej o nazwie *Datastore1\_on\_Openfiler\_Fiber\_Channel*. Utworzono go dla każdego z serwerów *ESXi* alokowanych w pamięci masowej typu *SAN Fibre Channel*.

### 3. System Radiografii Pośredniej *Synapse* w pamięciach masowych

System Radiografii Pośredniej *Synapse* jest systemem medycznym f-my *FUJIFILM Medical Systems USA, Inc.* Używany jest w Szpitalu Specjalistycznym im. Prof. E. Michałowskiego *MEDHOLDING S.A.* i służy do gromadzenia oraz prezentacji rentgenogramów (RTG) i tomogramów (TK). Eksploatowany jest na serwerze *ESXi* z dyskami wirtualnymi zlokalizowanymi w jego zasobach i posiada wewnętrzną strukturą przedstawioną na Rys. 1. Instalację systemu *Synapse* w pamięciach masowych, współpracujących z klastrem *ESXi* w sieci laboratoryjnej, przeprowadzono za pomocą metody kopiowania dysków wirtualnych [76], za pomocą polecenia *scp* wbudowanego w system *ESXi*. W ten sposób przeniesiono

dyski wirtualne ze środowiska produkcyjnego do środowiska laboratoryjnego w dwóch etapach. W etapie pierwszym przekopiowano, odpowiednie pliki (\*.vmdk) z maszyny wirtualnej *ESXi* hosta produkcyjnego do pamięci masowej klastra laboratoryjnego. W etapie drugim, na laboratoryjnym hoście *ESXi*, utworzono nową maszynę wirtualną (bez dysku) do której, jako dyski twarde, podłączono skopiowane pliki (\*.vmdk).

### 3.1. Instalacja systemu *Synapse* w pamięci masowej

Wykorzystując ww. metodę kopiowania dysków wirtualnych, w środowisku laboratoryjnym odtworzono w pamięci masowej typu *Fiber Channel SAN Storage*, pracującej pod kontrolą sytemu *Linux Openfiler ESA*, działającą kopię systemu systemu Radiografii Pośredniej *Synapse*, co pokazano na Rys. 3. Wykazano w ten sposób możliwość uruchomienia fragmentu specjalistycznego szpitalnego systemu *Synapse* przy wykorzystaniu pamięci masowej typu *block-based Fiber Channel SAN Storage* oraz potwierdzono przydatność opisywanej pamięci masowej, jako realnej alternatywy do migracji tego systemu na inne systemy pamięci masowych.

W podobny sposób odtworzono system *Synapse* w pamięci masowej typu *block-based Lenoovo Storage S3200*, *iSCSI* oraz w przestrzeni dyskowej typu *file-based* udostępnionej za pomocą protokołu *NFS*.

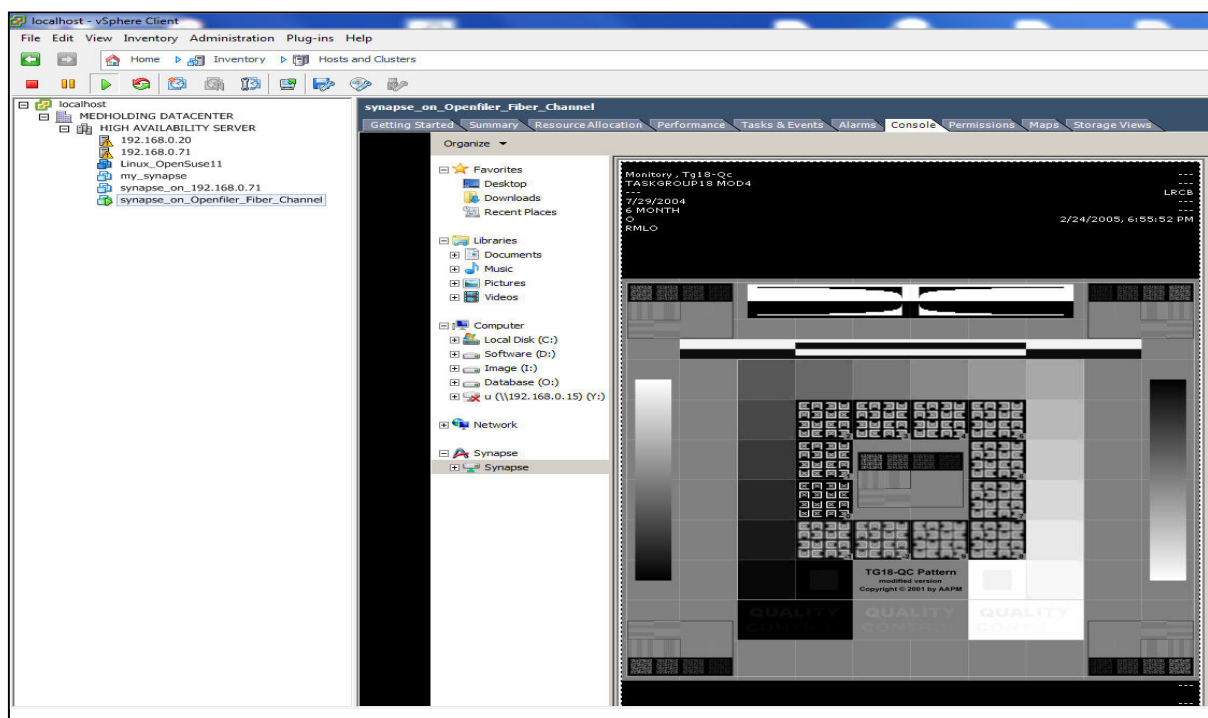
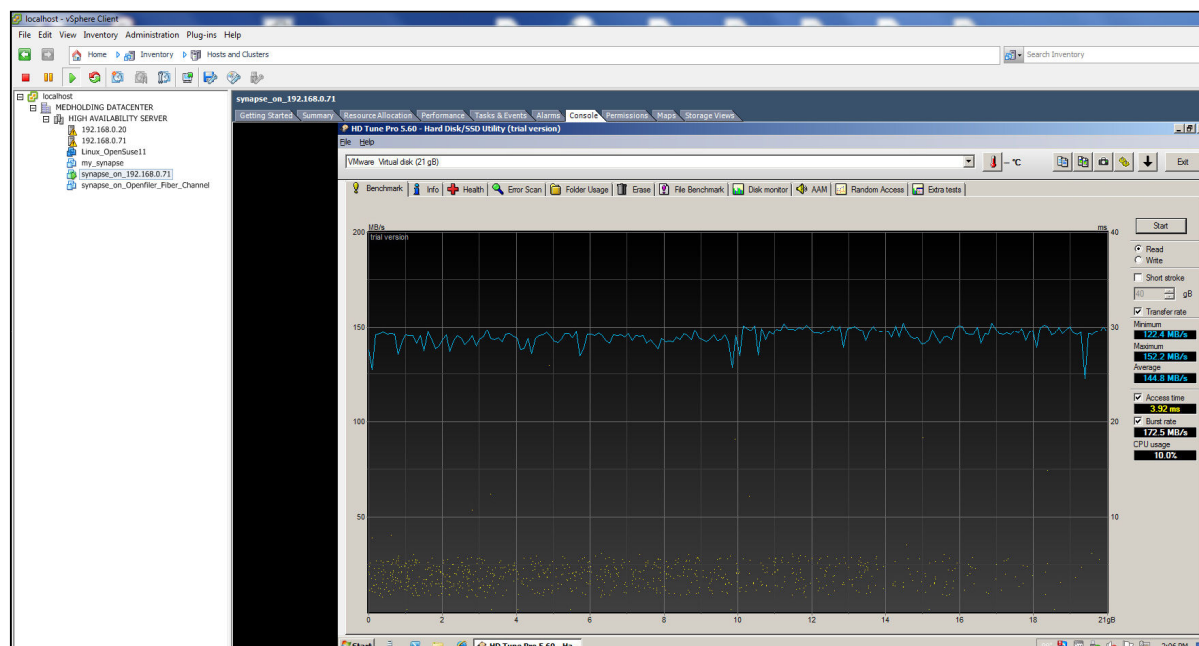


Fig. 3. System *Synapse* in mass *Fiber Channel SAN Storage* – image Test



### 3.2. Ocena porównawcza wirtualnych dysków systemowych z systemu Synapse

Ocenę porównawczą szybkości działania Systemu Radiografii Pośredniej *Synapse* dokonano za pomocą programu typu *Benchmark* o nazwie *HD Tune Pro* f-my *EFD Software*, służącego do testowania dysków poprzez pomiar prędkości odczytu dysków. Na rys. 4. pokazano wykres prędkości odczytu systemowego dysku wirtualnego *Synapse*, umiejscowionego w jednym z serwerów klastra *ESXi*, dokonany za pomocą tego programu.



Rys. 4. Prędkość odczytu systemowego dyska wirtualnego systemu *Synapse* umiejscowionego w jednym z serwerów klastra *ESXi*. Przeciętna prędkość odczytu: 144.8 MB/s  
Fig. 4. The speed of read *Synapse* virtual system disk located on one of the *ESXi* cluster's server. Average read speed: 144.8 MB/s

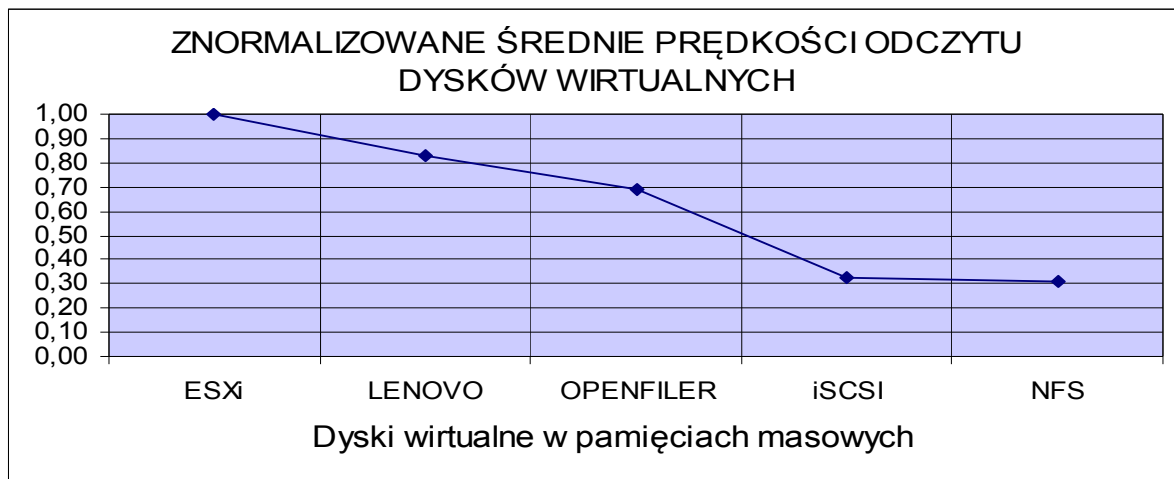
Ocenę porównawczą szybkości działania systemu *Synapse* wykonano dla kilku pamięci masowych włączonych do sieci laboratoryjnej, na których zainstalowano wirtualne maszyny *Synapse* wraz z wirtualnymi dyskami zawierającymi dyski systemowe z bazą danych *Oracle* przechowującą rentgenogramy i tomogramy, alokowane w:

1. pamięci masowej *block-based* jednego z serwerów *ESXi* klastra *HA*,
2. pamięci masowej *block-based* *Lenovo Storage S3200* włączonej do sieci *SAN* (4GB),
3. pamięci masowej *block-based* *Fiber Channel SAN Storage*, pracującej pod kontrolą *Openfiler ESA*, włączonej do sieci *SAN* (4GB),
4. pamięci masowej *block-based* *iSCSI* serwera w laboratoryjnej sieci *IP* (1GB),
5. systemie plików *file-based* *NFS* serwera w laboratoryjnej sieci *IP* (1GB).

Zmierzone średnie prędkości odczytu dysku wirtualnego zawierającego dyski systemowe z systemu Radiografii Pośredniej *Synapse* wyniosły:

1. 144,8 MB/s - z dysku wirtualnego alokowanego w pamięci masowej typu *block-based* w jednym z serwerów *ESXi* klastra *HA*,
2. 120,0 MB/s - z dysku wirtualnego w pamięci masowej *Lenovo Storage S3200*,
3. 100,4 MB/s - z dysku wirtualnego w pamięci masowej *Openfiler ESA*,
4. 47,3 MB/s - z dysku wirtualnego w pamięci masowej *iSCSI*,
5. 44,8 MB/s - z dysku wirtualnego w pamięci masowej systemu plików *NFS*.

Na Rys. 5 pokazano znormalizowane prędkości odczytu dla dysków wirtualnych alokowanych badanych pamięciach masowych, odniesionych do prędkości odczytu dysku systemowego alokowanego w jednym z serwerów *ESXi* klastra *HA*.



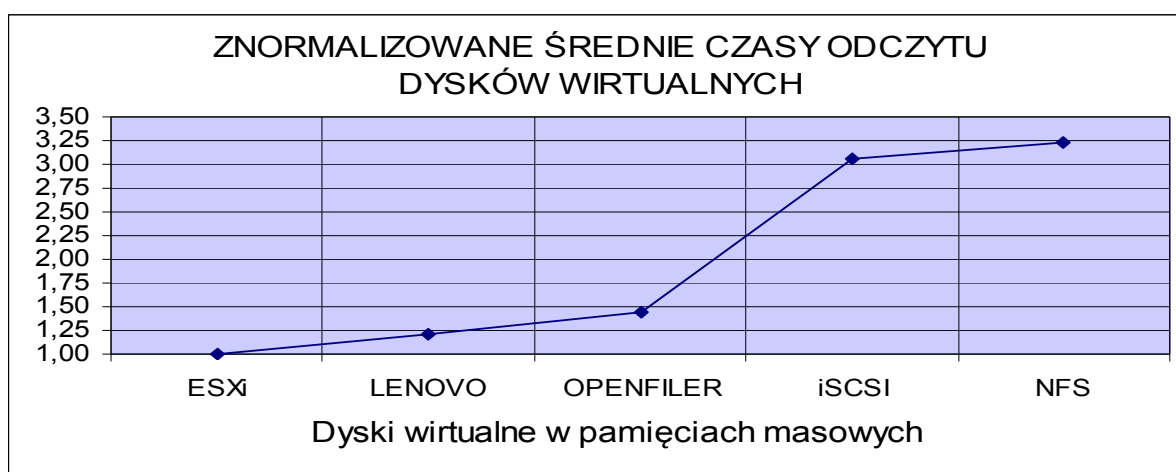
Rys. 5. Znormalizowane prędkości odczytu wirtualnych dysków w pamięciach masowych  
Fig. 5. Normalized read speed virtual disks in the mass storages

Obserwacje wynikające z Rys. 5 są następujące:

1. Znormalizowane prędkości odczytu wirtualnych dysków alokowanych w pamięciach masowych *SAN* i *NAS* są mniejsze od prędkości odczytu wirtualnych dysków alokowanych w pamięci masowej typu *block-level*, alokowanej na dyskach lokalnych jednego z serwerów klastra *ESXi* udostępnionej za pomocą protokołu *SCSI*.
  - 1.1 Procentowe prędkości odczytu, w stosunku do w pamięci masowej typu *block-level*, alokowanej na dyskach lokalnych jednego z serwerów klastra *ESXi*, dla tych dla tych pamięci wynoszą:
    - 70-85% dla pamięci *SAN* – pamięć masowa *LENOVO* oraz *OPENFILER*,
    - 30% dla pamięci *NAS* - pamięć masowa *iSCSI* oraz *NFS*.
2. Znormalizowane prędkości odczytu wirtualnych dysków alokowanych w pamięciach masowych *NAS*, włączonych do laboratoryjnej sieci *Ethernet*, alokowane w pamięci masowej typu:
  - *block-level iSCSI*, udostępniane za pomocą protokołu *FCoE*,

- *file-level* z systemem plików *NFS*, udostępniane za pomocą protokołu *NFS*, różnią się nieznacznie od siebie. Prędkość odczytu z pamięci masowej typu *block-level* *iSCSI* jest nieco większa niż prędkość odczytu z pamięci masowej typu *file-level* z systemem plików *NFS*. Ta obserwacja jest zgodna, co do tendencji, z obserwacjami podanymi w pracy [39].

Na Rys. 6 pokazano znormalizowane czasy odczytu dla dysków wirtualnych alokowanych badanych pamięciach masowych, odniesionych do czasu odczytu dysku systemowego alokowanego w jednym z serwerów *ESXi* klastra *HA*.



Rys. 6. Znormalizowane czasy odczytu wirtualnych dysków w pamięciach masowych  
Fig. 6. Normalized time read virtual disks in the mass storages

Obserwacje wynikające z Rys. 6 są następujące:

1. Znormalizowane czasy odczytu wirtualnych dysków alokowanych w pamięciach masowych *SAN* oraz *NAS* są większe od czasu odczytu wirtualnych dysków alokowanych w pamięci masowej typu *block-level*, alokowanej na dyskach lokalnych jednego z serwerów klastra *ESXi* udostępnionej za pomocą protokołu *SCSI*.
  - 1.1 Procentowe czasy odczytu, w stosunku do pamięci masowej typu *block-level*, alokowanej na dyskach lokalnych jednego z serwerów klastra *ESXi*, dla tych pamięci wynoszą:
    - 125-150% dla pamięci *SAN* – pamięć masowa *LENOVO* oraz *OPENFILER*,
    - 300-325% dla pamięci *NAS* - pamięć masowa *iSCSI* oraz *NFS*.
2. Znormalizowane czasy odczytu wirtualnych dysków zawierającymi dyski systemowe z bazą danych *Oracle* alokowanych w pamięciach masowych *NAS*, włączonych do laboratoryjnej sieci *Ethernet*, alokowane w pamięci masowej typu:
  - *block-level* *iSCSI*, udostępniane za pomocą protokołu *FCoE*,
  - *file-level* z systemem plików *NFS*, udostępniane za pomocą protokołu *NFS*,

różnią się nieznacznie od siebie. Czas odczytu z pamięci masowej typu *block-level iSCSI* jest nieco mniejszy niż czas odczytu z pamięci masowej typu *file-level* z systemem plików *NFS*. Ta obserwacja jest zgodna, co do tendencji, z obserwacją podaną w pracy [40] analizującej wpływ środowiska użytego do konfiguracji systemu bazodanowego na jego efektywność i wydajność, odniesioną do *Oracle*'owej bazy danych.

### 3.3. Ocena porównawcza heterogenicznych dysków wirtualnych Systemu Radiografii Pośredniej *Synapse*

W celu dokonania oceny porównawczej heterogenicznych dysków wirtualnych systemu *Synapse*, które rozpostarto w zbiorze heterogenicznych pamięci masowych *Openfiler ESA*, *Lenovo Storage S3200*, *iSCSI*, *NFS*, w badanych pamięciach masowych instalowano kolejne maszyny wirtualne tego systemu, a następnie tworzono nowe dyski wirtualne w nowych obszarach typu *Datastore*, w kolejnych pamięciach masowych. Następnie, z poziomu systemu operacyjnego wirtualnej maszyny, istniejący dysk wirtualny rozszerzano na nowo utworzone wirtualne dyski. Rozszerzenia dysku wirtualnego dokonywano za pomocą narzędzia administracyjnego „*Computer Management*”, tj. programu „*Disk Management*”.

Zmierzone średnie prędkości odczytu heterogenicznych dysków wirtualnych, rozpostartych w zbiorach pamięci masowych, oznaczonych dalej w tekście jako: *Lenovo Storage*, *Openfiler ESA*, *iSCSI* oraz *NFS*, pokazano w Tab. 2.

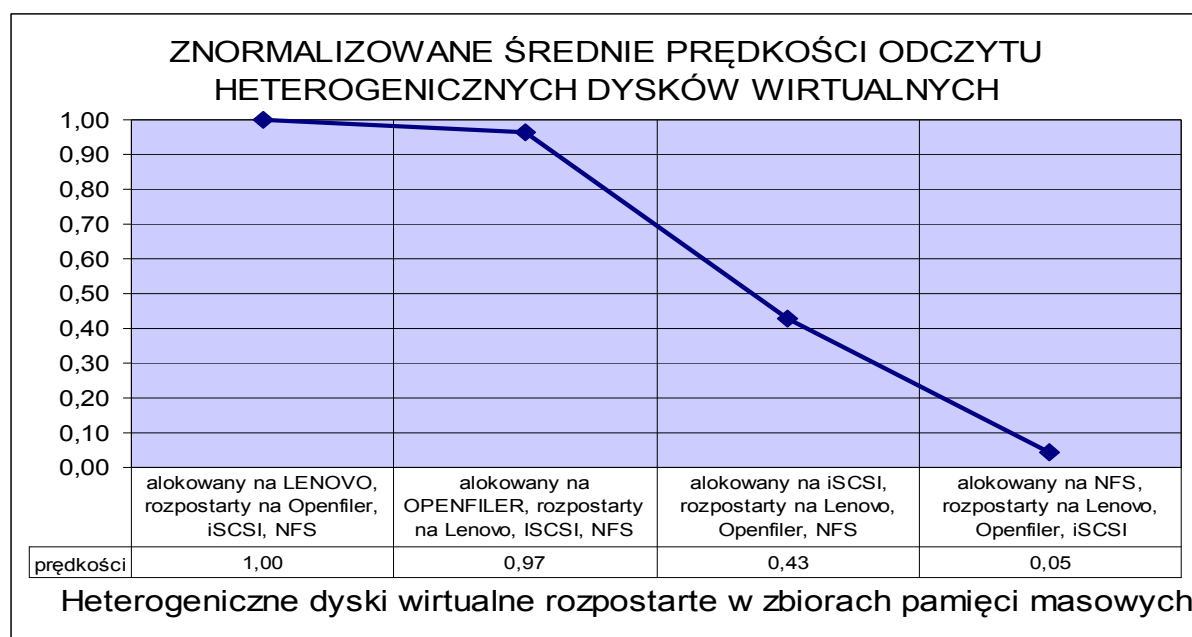
Prędkość odczytu [MB/s]	z heterogenicznego dysku wirtualnego alokowanego w pamięci masowej typu:	rozpostartego w zbiorze pamięci masowych:
115,5	Lenovo Storage	Openfiler ESA, iSCSI, NFS
111,5	Openfiler ESA	Lenovo, iSCSI, NFS
49,4	iSCSI	Lenovo, Openfiler, NFS
5,2	NFS	Lenovo, Openfiler, iSCSI

Tab. 2. Średnie prędkości odczytu heterogenicznych dysków wirtualnych

Tab 2. Average read speed of heterogeneous virtual disks

Ocenę porównawczą heterogenicznych dysków wirtualnych dokonano poprzez porównanie średniej prędkości ich odczytu w stosunku do średniej prędkości odczytu heterogenicznego dysku wirtualnego najszybszego z nich.

Na Rys. 7 pokazano znormalizowane prędkości odczytu, w stosunku do pamięci masowej *Lenovo*, dla heterogenicznych dysków wirtualnych alokowanych w pamięciach masowych *Lenovo*, *Openfiler ESA*, *iSCSI*, *NFS* i rozpostartych w zbiorze kolejnych badanych pamięciach masowych.



Rys. 7. Znormalizowane średnie prędkości odczytu w heterogenicznych dyskach wirtualnych alokowanych w pamięciach masowych

Fig. 7. Normalized average read speed in heterogeneous virtual disks allocated in mass storages

Obserwacje wynikające z Rys. 7 są następujące:

1. Znormalizowane prędkości odczytu heterogenicznych wirtualnych dysków alokowanych w pamięciach masowych *SAN* typu *block-level*, tj. LENOVO oraz OPENFILER i rozpostartych na pozostałych typach pamięci masowych różnią się nieznacznie,
2. Znormalizowana prędkość odczytu heterogenicznego wirtualnego dysku alokowanego w pamięciach masowych *NAS*, tj.:
  - iSCSI typu *block-level*, rozpostartego na pozostałych typach pamięci masowych jest znacznie niższa niż prędkości odczytu heterogenicznych wirtualnych dysków alokowanych w pamięciach masowych *SAN*,
  - NFS typu *file-level*, rozpostartego na pozostałych typach pamięci masowych jest nieporównywalnie znacznie niższa niż prędkości odczytu heterogenicznych wirtualnych dysków alokowanych w pamięciach masowych *SAN*,
3. Znormalizowana prędkość odczytu heterogenicznego wirtualnego dysku alokowanego w pamięci masowej *iSCSI* typu *block-level*, rozpostartego na pozostałych typach pamięci masowych jest wyraźnie większa niż znormalizowana prędkość odczytu heterogenicznego wirtualnego dysku alokowanego w pamięci masowej *NFS* typu *file-level*, alokowanego na pozostałych typach pamięci masowych.

Kolejną ocenę porównawczą heterogenicznych dysków wirtualnych dokonano poprzez porównanie średnich czasów odczytu w stosunku do średniego czasu odczytu heterogenicznego dysku wirtualnego najszybszego z nich. Na Rys. 8 pokazano znormalizowane czasy odczytu, w stosunku do pamięci masowej Lenovo, dla heterogenicznych dysków wirtualnych alokowanych w pamięciach masowych Lenovo, Openfiler ESA, iSCSI, NFS i rozpostartych w zbiorze kolejnych badanych pamięciach masowych.



Rys. 8. Znormalizowane średnie czasy odczytu w heterogenicznych dyskach wirtualnych alokowanych w pamięciach masowych

Fig. 8. Normalized average read time in heterogeneous virtual disks allocated in mass storages

Obserwacje wynikające z Rys. 8 są następujące:

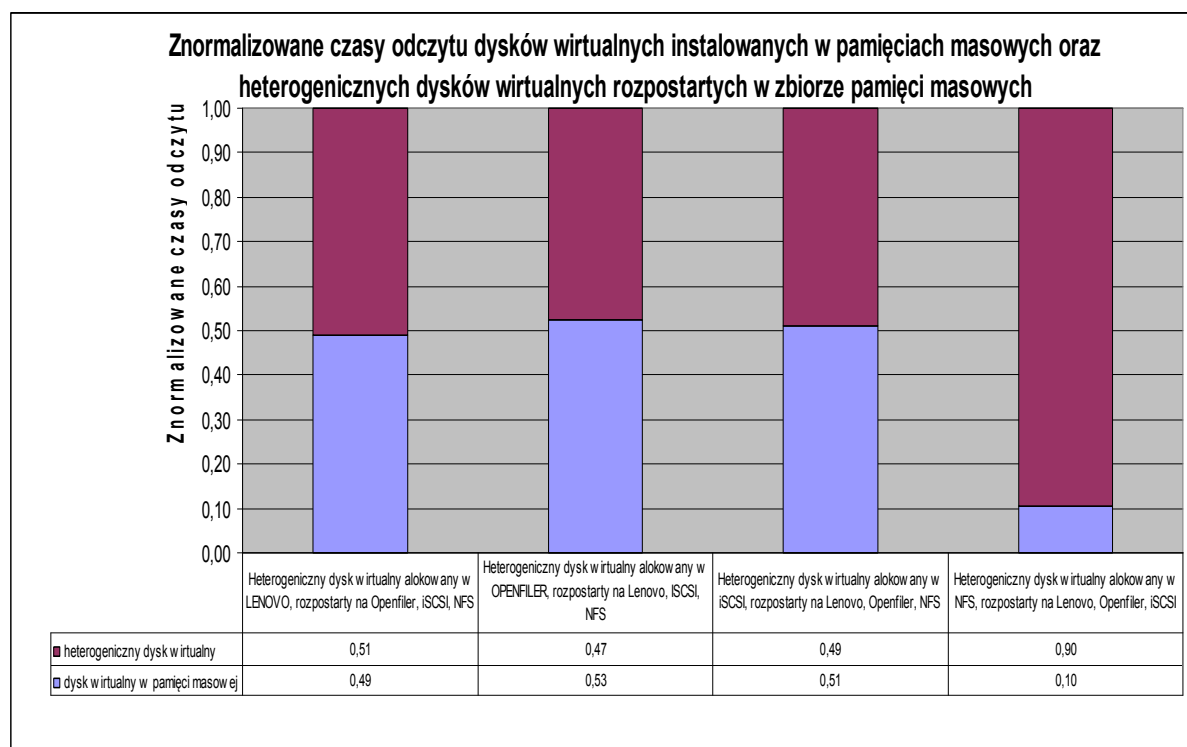
1. Znormalizowane czasy odczytu heterogenicznych wirtualnych dysków alokowanych w pamięciach masowych SAN typu *block-level*, tj. LENOVO oraz OPENFILER i rozpostartych na pozostałych typach pamięci masowych są porównywalne,
2. Znormalizowany czas odczytu heterogenicznego wirtualnego dysku alokowanego w pamięciach masowych NAS, tj.:
  - iSCSI typu *block-level*, rozpostartego na pozostałych typach pamięci masowych jest dwukrotnie większy w stosunku do średniego czasu odczytu heterogenicznych wirtualnych dysków alokowanych w pamięciach masowych SAN,

- NFS typu *file-level*, rozpostartego na pozostałych typach pamięci masowych jest kilkunastokrotnie większy od czasu odczytu heterogenicznych wirtualnych dysków alokowanych w pamięciach masowych *SAN*,
3. Znormalizowany czas odczytu heterogenicznego wirtualnego dysku alokowanego w pamięci masowej iSCSI typu *block-level*, rozpostartego na pozostałych typach pamięci masowych jest dziesięciokrotnie krótszy niż znormalizowany czas odczytu heterogenicznego wirtualnego dysku alokowanego w pamięci masowej NFS typu *file-level*, alokowanego na pozostałych typach pamięci masowych.

Kończąc dyskusję dotyczącą oceny porównawczej heterogenicznych dysków wirtualnych, na Rys. 9 przedstawiono porównanie znormalizowanych czasów odczytu dysków wirtualnych instalowanych w pamięciach masowych Lenovo, Openfiler ESA, iSCSI, NFS oraz znormalizowanych czasów odczytu tych dysków instalowanych w tych pamięciach i rozszerzonych do postaci heterogenicznych dysków wirtualnych i rozpostartych w zbiorze kolejnych badanych pamięci masowych. Znormalizowane czasy odczytu, pokazane w Tab. 3, stanowią czasy względne, obliczone jako stosunek bezwzględnego czasu odczytu konkretnego dysku wirtualnego do sumy odczytów czasów bezwzględnych dysku wirtualnego i heterogenicznego dysku wirtualnego, stanowiące razem 100%.

Dysk wirtualny w pamięci masowej:	Średni czas odczytu dysku wirtualnego [ms]	Heterogeniczny dysk wirtualny alokowany w:	Średni czas odczytu heterogenicznego dysku wirtualnego [ms]	Znormalizowany średni czas odczytu dysku wirtualnego (po normalizacji do 100%)	Znormalizowany średni czas odczytu heterogenicznego dysku wirtualnego (po normalizacji do 100%)
LENOVO	8,33	LENOVO, rozpostarty na Openfiler, iSCSI, NFS	8,66	<b>0,49</b>	<b>0,51</b>
OPENFILE R	9,96	OPENFILER, rozpostarty na Lenovo, iSCSI, NFS	8,97	<b>0,53</b>	<b>0,47</b>
iSCSI	21,14	iSCSI, rozpostarty na Lenovo, Openfiler, NFS	20,24	<b>0,51</b>	<b>0,49</b>
NFS	22,32	NFS, rozpostarty na Lenovo, Openfiler, iSCSI	192,31	<b>0,10</b>	<b>0,90</b>

Tab 3. Average and normalized read time of heterogeneous virtual disks



Rys. 9. Znormalizowane względne średnie czasy odczytu w dyskach wirtualnych  
Fig. 9. Normalized relative average read time in virtual disks

Obserwacje wynikające z Rys. 9 są następujące:

1. Znormalizowane względne czasy odczytu dysków wirtualnych instalowanych w pamięciach masowych typu *block-level*, tj.: LENOVO, OPENFILER, iSCSI porównane do heterogenicznych dysków wirtualnych alokowanych w tych pamięciach i rozpostartych w zbiorze pamięci masowych typu *block-level* ora *file-level* mają podobne względne czasy odczytu.
2. Znormalizowany względny czas odczytu heterogenicznego dysku wirtualnego alokowanego w pamięci masowej NFS typu *file-level* i rozpostartego w zbiorze pamięci masowych typu *block-level*, tj.: LENOVO, OPENFILER, iSCSI jest ok. dziewięciokrotnie większy od czasu odczytu dysku wirtualnego instalowanego w pamięci masowej NFS.

#### 4. Podsumowanie, wnioski i propozycje

W niniejszej pracy wykazano możliwości uruchomienia fragmentu szpitalnego Systemu Radiografii Pośredniej *Synapse* w pamięci masowej typu *Fiber Channel SAN Storage*, działającej pod kontrolą systemu operacyjnego *Linux Openfiler ESA* zapewniającego obsługę



blokowych pamięci masowych, współpracującą z klastrem wysokiej dostępności *HA VMWARE ESXi*.

Ponadto, na gruncie przyjętej koncepcji heterogenicznych woluminów dysków wirtualnych, w pracy wykazano:

1. możliwość rozszerzania dysku wirtualnego systemu *Synapse*, przechowującego rentgenogramy i tomogramy, do postaci heterogenicznego dysku wirtualnego rozpostartego w zbiorze heterogenicznych blokowych pamięci masowych *SAN* oraz plikowych pamięci masowych *NAS*,
2. rozszerzanie dysku wirtualnego do postaci heterogenicznego dysku wirtualnego zapewnia prawidłowe funkcjonowanie tego systemu, przy jednoczesnym radykalnym skróceniu czasu przestoju systemu informatycznego,
3. faktyczny czas przestoju systemu *Synapse* został radykalnie zredukowany do czasu potrzebnego na rozszerzanie jego dysku wirtualnego do postaci heterogenicznej.

#### Wnioski – część 1

1. Gdy System Radiografii Pośredniej *Synapse* działa w pamięci masowej typu *block-level* na dyskach lokalnych jednego z serwerów *ESXi* klastra włączonego do sieci laboratoryjnej (1GB), udostępniającego dane stacji roboczej za pomocą protokołu *FCoE (Fibre Channel over Ethernet)*, to czas wyszukiwania i wyświetlania rentgenogramów na stacji roboczej dla konkretnych pacjentów jest najkrótszy; warunki **bardzo dobre** do normalnej diagnostyki lekarskiej.
2. W przypadku, gdy system *Synapse* działa w pamięciach masowych *SAN* typu *block-level Fiber Channel SAN Storage* działającej pod kontrolą systemu *Openfiler ESA* lub *LENOVO STORAGE S3200*, wówczas szybkość uzyskiwania poszczególnych rentgenogramów jest średnia i porównywalna dla obydwu pamięci; warunki **dobre** do normalnej diagnostyki lekarskiej.
3. W sytuacji, kiedy system *Synapse* działa w pamięci masowej typu *file-level NFS* lub typu *block-level iSCSI* serwera linuxowego włączonego do laboratoryjnej sieci Ethernet, to szybkość uzyskiwania poszczególnych rentgenogramów jest najmniejsza; warunki **akceptowalne** do normalnej diagnostyki lekarskiej. Zwiększenie wydajności plikowych pamięci masowych *NFS* lub blokowych *iSCSI* można poprawić poprzez zwiększenie przepustowości sieci i parametrów macierzy dyskowej, na której są one zainstalowane.

#### *Propozycja 1.*

Ponieważ w sieci laboratoryjnej z Rys. 1 użyto przełącznika *SAN FC 200E* o przepustowości 4Gbps, zwiększenie szybkości Systemu Radiografii Pośredniej *Synapse* działającego w

pamięci masowej typu *Fiber Channel SAN Storage* pod kontrolą systemu operacyjnego *Openfiler ESA* lub też w pamięci masowej *LENOVO STORAGE S3200* można osiągnąć poprzez zastosowanie szybszego przełącznika w sieci *SAN*, np. *Bracode 300* (8 Gbps) [77], *Bracode G620* (32-Gbps) [78] lub *Brocade Gen 6 platform* (32-Gbps) [79].

### Wnioski – część 2

1. W przypadku, gdy dysk wirtualny systemu *Synapse*, przechowujący rentgenogramy i tomogramy, alokowany jest w heterogenicznej pamięci masowej typu *block-level*:
  - LENOVO, rozpostartej w zbiorze pamięci masowych *Openfiler*, *iSCSI*, *NFS*,
  - pracującej pod kontrolą systemu operacyjnego *Openfiler ESA*, rozpostartej w zbiorze pamięci masowych *Lenovo Storage S2300*, *iSCSI* oraz *NFS*,wówczas szybkość uzyskiwania poszczególnych rentgenogramów jest średnia i porównywalna dla obydwu pamięci; warunki **dobrze** do normalnej diagnostyki lekarskiej.
2. W przypadku, gdy dysk wirtualny systemu *Synapse*, przechowujący rentgenogramy i tomogramy, alokowany jest w heterogenicznej pamięci masowej *NFS* typu *file-level*, rozpostarty w zbiorze pamięci masowych typu *block-level*, tj.: *Lenovo Storage S2300*, *iSCSI*, wówczas szybkość uzyskiwania poszczególnych rentgenogramów jest najmniejsza; warunki **nie akceptowalne** do normalnej diagnostyki lekarskiej.

### LITERATURA

1. Fujifilm Medical Systems, Foundation Technologies, *Synapse Release Version 4.0.0*, MKT-0023291-A, 2012.
2. Fujifilm *Synapse Deployment Guide*, KEMP Technologies, Inc., 2015.
3. *vSphere ESXi Bare-Metal Hypervisor*, VMware Polska, 2015.
4. Daniluk D.: *Wirtualne serwery na bazie oprogramowania VMware GSX/ESX wspomaganego przez VMware VirtualCenter*, 2006.
5. *PowerEdge T420 Technical Guide*, Dell Inc., 2014.
6. *Configure RAID on a Dell Server*, vSphere 5 Documentation Center, [https://pubs.vmware.com/vsphere-50/index.jsp?topic=%2Fcom.vmware.vsa.doc\\_10%2FGUID-36314EB3-0FC6-4172-B757-71BEA66236F3.html](https://pubs.vmware.com/vsphere-50/index.jsp?topic=%2Fcom.vmware.vsa.doc_10%2FGUID-36314EB3-0FC6-4172-B757-71BEA66236F3.html).
7. *Fibre Channel SAN Configuration Guide – VMware*, EN-000109-05, 2010.
8. *QNAP in vSphere Environment*, QNAP Systems, Inc., 2009.
9. *exacqVision Virtual Appliance User Manual*, 2017.
10. Krenn T.: *Zwiększenie pojemności wirtualnego dysku w VMware*, 2012.
11. *Install Dell OpenManage Server Administrator on VMware ESXi 5.x*, 2014.
12. *Dell OpenManage Server Administrator (OMSA) 8.2 Install for ESXi 6.0*, 2017.

13. VMware vSphere 5.5 Documentation, VMware Inc., 2017, <https://pubs.vmware.com/vsphere-55/index.jsp>.
14. Fibre Channel SAN Configuration Guide – VMware, EN-000109-05, 2010.
15. VMware vSphere 5.1 Documentation Center, <https://pubs.vmware.com/vsphere-51/index.jsp#com.vmware.vsphere.doc/GUID-1B959D6B-41CA-4E23-A7DB-E9165D5A0E80.html>.
16. Mon-Chau Shiu: Storage Area Network, Infortrend Technology, 2007.
17. Smith E., Hultman R., Klopping D.: Fibre Channel SAN Topologies, EMC Corp., 2015.
18. Pico J.: An Analysis From A Security Perspective, Fibre Channel Storage Area Networks - SANS Institute, 2006.
19. Norman D.: Fibre Channel Technology for Storage Area Networks, 2001.
20. Black D.L.: A Storage Menagerie: SANs, Fibre Channel, Replication and Networks, 2011.
21. Dale D.: iSCSI & Fibre Channel SANs, Storage Networking Industry Association, 2007.
22. Fibre Channel Storage area Network, <https://www.google.pl/url?sa=t&rct=j&q=&esrc=s&source=web&cd=73&ved=0ahUKEwiQ-6me4trSAhVFVBQKHdj8AIE4RhAWCCgwAg&url=http%3A%2F%2Fjbiet.edu.in%2Fcoursefiles%2Fcse%2FHO%2Fcse4%2FSAN4.pdf&usg=AFQjCNGuJ-qgJaBf9f1Kt-hAfMhE-HxFyw&sig2=5iJgB9TpakoLLkGcM11-Sw&cad=rja>.
23. VMware vSphere 5.1: 16Gb Fibre Channel SANs with HP ProLiant DL380 Gen8 servers and HP 3PAR Storage, Deployment Guide, 2013.
24. Openfiler, <https://www.openfiler.com/products>.
25. Bastiaansen R.: Install and configure Openfiler for ESXi shared storage with NFS and iSCSI, 2013.
26. Seget V.: How to configure OpenFiler iSCSI Storage for use with VMware ESX, 2015.
27. Openfiler Administration Guide, Openfiler Ltd. UK, 2009.
28. Tivari R.: How to install openfiler 2.99, 2013.
29. Intel: Delivering Low Cost High IOPS VM Datastores Using Openfiler, Document Number: 329238-002US, 2014.
30. How to configure Openfiler v2.3 iSCSI Storage with VMware ESXi v4, 2017.
31. Introduction to a “Network File System” (NFS), 2009.
32. Shepler S., Callaghan B., Robinson D., Thurlow R., Beame C., Eisler M., Noveck D.: Network File System (NFS) version 4 Protocol, RFC 3530, IETF, DOI: 10.17487/RFC3530, OCLC 943595667, 2003.
33. Callaghan B.: Pawlowski B., Staubach P.: NFS Version 3 Protocol Specification, RFC 1813, IETF, DOI: 10.17487/RFC1813, OCLC 943595667, 1995.
34. Nowicki B.: NFS: Network File System Protocol specification, RFC 1094, IETF, DOI: 10.17487/RFC1094, OCLC 943595667, 1989.
35. iSCSI SAN Configuration Guide - VMware, VMware, Inc., Revision: 20090313, Item: EN-000035-01, 2009.
36. iSCSI SAN Configuration Guide - ESX 4.1 - VMware, VMware, Inc., Revision: EN-000288-04, 2010.

37. Sudhish P. T.: Multipathing Configuration for Software iSCSI Using Port Binding - VMware, Technical white paper, VMware Inc., 2012.
38. Configuring VMware vSphere Software iSCSI With Dell EqualLogic PS Series Storage, Dell technical report, TR1049 V1.2, 2010.
39. Radkov P., Yin L., Goyal P., Sarkar P., Shenoy P.: An Experimental Comparison of Block and File Access Protocols for IP Networked Storage, Department of Computer Science, University of Massachusetts.
40. Hareźlak K., Werner A., Bach M., Duszeńko A.: The Storage Organisation Influence on Database Operations Performance, 2014.
41. Bandulet C.: The storage evolution: From blocks, files and blocks, to object storage system, Storage Networking Industry Association, 2007.
42. Kazuo G., Kitsuregawa M.: The History of Storage Systems, Proceedings of the IEEE, Vo. 100, 2012.
43. Pillai S.: SAN vs NAS - Difference between a Storage Area Network and Network Attached Storage, 2014.
44. Yang J.: Network Attached Storage, 2015.
45. Pradnya E. G., Sachin B.: A Survey Data Storage Technologies, International Journal of Engineering Science and Innovative Technology (IJESIT) Volume 2, Issue 2, March 2013.
46. Heng L.: Storage Area Network Architectures, Technology White Paper, PMC-2022178, Issue 1: 2003.
47. Khattar R., Murphy S., Tarella G., Nystrom K.: Storage Area Network SAN, IBM International Technical Support Organization, 1999.
48. Riabov V.: Storage Area Networks Fundamentals, 2007.
49. Yang J.: Network Attached Storage, 2015.
50. Block-Level Backups, Commvault Systems Inc., 2017.
51. Hogan C.: Storage Protocol Comparison White Paper, Technical Marketing Documentation v 1.0, 2012.
52. Onisick J.: Storage Protocols, 2010.
53. Traeger A.: An Introduction to NFS. Network-Based File Systems, IBM Haifa Research Lab Internal Storage Course v1.0, 2012.
54. SCSI Architecture Model”, ANSI X3.270, 1996.
55. SCSI Architecture Model - 3 (SAM-3), Revision -1, 2002.
56. Dooley R.: Fundamentals of SCSI and iSCSI, 2011.
57. iSCSI Protocol Concepts and Implementation, White paper, Cisco Systems, 2001.
58. iSCSI (RFC 3720), Internet Engineering Task Force, 2004.
59. Jonghoon J.: iSCSI SAN Topologies, EMC Corporation, 2015.
60. Burton R.: Fibre Channel, 2000.
61. Norman D.: Fibre Channel Technology for Storage Area Networks, 2001.
62. Fibre Channel, INCITS xxx-200x Fibre Channel Backbone - 5 Rev 2.00, T11/Project, 1871-D/Rev 2.00, 2009.
63. Tate J., Beck P., Ibarra H., Kumaravel S., Miklas L.: Introduction to Storage Area Networks, 2016.
64. FCoE Initialization Protocol (FIP), Brocade education solutions, Revision 1009.
65. Lippitt M., Smith E., Hughes D.: Fibre Channel over Ethernet (FCoE) Data Center Bridging (DCB) Concepts and Protocols, EMC Corporation, 2015.
66. Platt T.: Compression in HFS+: Space savings and performance gain, 2009.

67. Block-Level Compression, Teradata Corporation, 2013.
68. Klonatos Y., Makatos T., Marazakis M., Flouris M. D.: Transparent Online Storage Compression at the Block-Level, Transparent online storage compression at the block-level. ACM Trans. Storage 8, 2, Article 5, 2012.
69. Makatos T., Klonatos Y., Marazakis M., Flouris M., Bilas A.: ZBD: Using Transparent Compression at the Block Level to Increase Storage Space Efficiency, Department of Computer Science, University of Crete.
70. Bożyk K.: Technologia SAN (Storage Area Networks)-Praca dyplomowa magisterska, Samodzielny Zakład Sieci Komputerowych Politechnika, Łódź 2005.
71. Norman D.: Fibre Channel Technology for Storage Area Networks, 2001.
72. Brocade Communications Systems, Inc.: EZSwitchSetup, Administrator's Guide, 2011.
73. Brocade Communications Systems, Inc.: Secure SAN Zoning, Best Practices.
74. Weith W.: Deploying VMware Virtual Infrastructure 3.5 in a Heterogeneous Server and Storage Environment Using a Brocade 8 Gbps FC SAN Infrastructure, Brocade Communications Systems, Inc., 2009.
75. Brocade Communications Systems, Inc.: Brocade Guide to Understanding Zoning, Document number: 53-0000213-01, 2002.
76. Brzeżek I.: VMWare ESXi przenoszenie maszyn wirtualnych, 2013.
77. Brocade 300 Switch, Data sheet, Brocade Communications Systems, Inc., 2017.
78. Brocade G620 Switch, Data sheet, Brocade Communications Systems, Inc., 2017.
79. Brocade Communications Systems, Inc.: Brocade EZSwitchSetup, Administrator's Guide, 2015.

Recenzent:

Wpłynęło do Redakcji 30 lipca 2017 r.

## Abstract

This article presents the concept of heterogeneous virtual disks, outstretched hands in the mass storage SAN and NAS type collection, work with virtual machines in VMWARE ESXi, on the example of the portion of the specialized hospital System of indirect Radiography "Synapse" FUJIFILM Medical Systems USA, Inc., which is used in Specialized Hospital named Prof. E. Michalowski MEDHOLDING S.A., for the collection and presentation of the rentgenograms and tomographs.

On the basis of this concept, it has been shown that it is possible to extend the virtual disk from "Synapse" system to form a heterogeneous virtual disk and to ensure the proper functioning of the system in a situation a steady supply of new rentgenograms and tomographs. Extend virtual disk to form a heterogeneous virtual disk provides the proper functioning of the system and does not cause long-term downtime of the system, thus allowing normal medical Diagnostics.