

Rozwiązania Obiektowe HCP

HITACHI
Inspire the Next



Hitachi Content Platform

Rozwiązanie Hitachi Content Platform (HCP) jest często wykorzystywane jako podstawa do budowy centralnego obiektowego magazynu danych przez różne przedsiębiorstwa i organizacje. Używane jest ono wówczas w wielu różnych scenariuszach i dla wielu różnych aplikacji jednocześnie. Wśród tych scenariuszy, najpopularniejsze to:

- budowa chmury prywatnej dla danych i świadczenie usług wewnątrz i na zewnątrz organizacji,
- archiwizacja danych, z różnych środowisk i aplikacji,
- przechowywanie i analiza logów z różnych środowisk i aplikacji (np. integracja ze Splunkiem),
- synchronizacja oraz współdzielenie plików wewnątrz oraz na zewnątrz organizacji (HCP Anywhere),
- budowa obiektowego środowiska plikowego, bez konieczności wykonywania lokalnego backupu danych (HCP Anywhere Edge),
- indeksowanie danych i metadanych oraz ich pełnotekstowe wyszukiwanie (Hitachi Content Intelligence),
- przetwarzanie danych i metadanych, ich wzbogacanie, transformacja, filtrowanie, oznaczanie (ETL, czyli mechanizm Extract Transform Load) z wielu źródeł danych oraz budowanie kolekcji indeksów dla danych i metadanych (Hitachi Content Intelligence),
- archiwizacja danych z HDFSa do HCP (np. archiwizacja danych, które chcemy, aby były przechowywane w dłuższym okresie czasu poza klasterem Hadoop),
- wykonywanie analiz biznesowych bezpośrednio w jeziorze danych (data lake) HCP, z wykorzystaniem Spark'a.

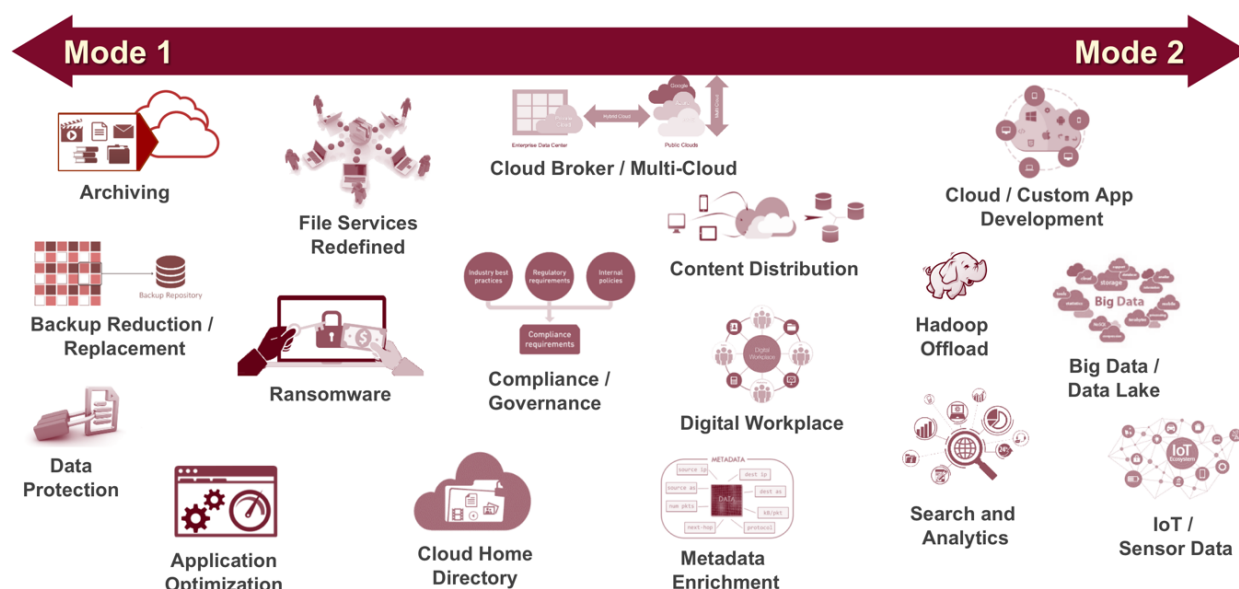
”

„Osiąganie tak dużych wydajności jest możliwe między innymi dzięki unikalnej architekturze urządzeń Hitachi NAS.”

Dzięki dostępnej w HCP funkcji multitenancy, istnieje możliwość konsolidacji tych wszystkich scenariuszy, projektów i aplikacji w ramach jednej macierzy obiektowej. To z kolei pozwala na uproszczenie zarządzania, obniżenie kosztów TCO oraz lepsze wykorzystanie przechowywanych w HCP danych. Mechanizm zarządzania metadanymi w HCP pozwala na oddzielenie danych od aplikacji i wykorzystanie ich zgodnie z oczekiwaniami biznesu (różnych linii biznesowych). Z kolei funkcjonalność logicznego partycjonowania HCP pozwala na zapewnienie i utrzymanie rozdzielności danych zapisywanych do HCP z różnych źródeł i umożliwia zarządzanie i wgląd do tych danych tylko przez określonych administratorów.

Przypadki wykorzystania HCP klasyfikowane są często w obszarze dwóch grup scenariuszy:

- **Mode 1** – tradycyjne, czyli takie, w których macierze obiektowe wykorzystywane są już od wielu lat
- **Mode 2** – nowoczesne, czyli takie, w których dopiero od niedawna zaczęto postrzegać wykorzystanie obiektowych magazynów danych jako dużą wartość dla biznesu.



Scenariusze wykorzystania Hitachi Content Platform.

Poniżej zaprezentowano kilka przykładów realizacji konkretnych celów biznesowych za pomocą określonych scenariuszy użycia Hitachi Content Platform.

Modernizacja strategii ochrony danych

1. Wdrożenie archiwizacji danych i przeniesienie danych, które nie są często używane, z systemów i aplikacji produkcyjnych do obiektowego magazynu danych (do archiwum on-line). Dane, które są zabezpieczone w archiwum nie wymagają backupu. Kopia zapasowa realizowana jest tylko dla tego zbioru danych, które pozostały w systemach produkcyjnych (dla danych aktywnych). Mniejsza ilość danych, które wymagają backupu oznacza: krótsze okna backupowe, szybszy czas wykonywania kopii oraz krótsze czasy odtworzenia w przypadku awarii.
2. Całkowite wyeliminowanie backupu oraz redukcja jego kosztów – przeniesienie dużych zbiorów danych plikowych (setki terabajtów

i petabajtów) z serwerów NAS, które wymagają nie tylko replikacji zdalnej, ale również lokalnego backupu do macierzy obiektowych z wbudowanymi mechanizmami WORM i z retencją danych. Wymienione funkcjonalności pozwalają na zapewnienie niezmienności oraz nieusuwalności danych przechowywanych w magazynie obiektowym. A co za tym idzie nie wymagają wykonywania kopii zapasowych danych przechowywanych wewnątrz. Zabezpieczamy się w tym przypadku jedynie na wypadek awarii całego ośrodka danych, korzystając z replikacji zdalnej (analogicznie jak w przypadku macierzy blokowych i plikowych).

3. Wykorzystanie obiektowych magazynów danych jako miejsce przechowywania kopii zapasowych (zamiast taśm). Nowoczesne aplikacje do backupu dostarczają konektorów do chmury (cloud connector). Korzystają one z bardzo wydajnych interfejsów dostępowych (S3, http) i pozwalają na wykonywanie kopii zapasowej danych do chmury, którą w szczególności może być również macierz obiektowa zainstalowana w centrum przetwarzania danych. Kopie zapasowe z retencją długoterminową mogą być wówczas przechowywane on-line (zamiast off-line) na dyskach macierzy obiektowej.
4. Ochrona przed złośliwym oprogramowaniem szyfrującym (ransomware). Dane przechowywane w obiektowych magazynach danych, wyposażonych w technologię WORM nie mogą zostać zmienione. Próba zaszyfrowania danych może zakończyć się albo utworzeniem kolejnej wersji obiektu (jeżeli wersjonowanie jest włączone), albo niepowodzeniem. W jednym i w drugim przypadku oryginalny dokument pozostaje nienaruszony.

Modernizacja usług i transformacja infrastruktury IT w kierunku środowiska usługowego

1. Budowa środowiska pracy cyfrowej. Dostęp do danych użytkowników z dowolnego miejsca i dowolnego urządzenia oraz w dowolnym czasie. Możliwość współdzielenia plików pomiędzy użytkownikami, wewnątrz i na zewnątrz organizacji oraz dodatkowo synchronizacja dokumentów pomiędzy różnymi urządzeniami (laptop, tablet, smartfon). Wszystkie dane zabezpieczone i przechowywane w obiektowym magazynie danych.
2. Zdalne wykorzystanie bramy plikowej jako urządzenia dostępowego do chmury prywatnej zbudowanej w centrum przetwarzania danych (do obiektowego magazynu danych). Automatyczna synchronizacja danych, realizowana w czasie rzeczywistym pomiędzy zdalnym urządzeniem dostępowym, a centralnym magazynem danych. Możliwość współdzielenia plików pomiędzy oddziałami.

Modernizacja zarządzania danymi w kontekście zgodności z regulacjami

1. Wykorzystanie technologii WORM, retencji danych oraz innych wbudowanych w macierz obiektową mechanizmów na potrzeby spełnienia wewnętrznych polityk przedsiębiorstwa oraz zewnętrznych regulacji prawnych. Zapewnienie zgodności nie tylko w samej aplikacji, ale również na poziomie sprzętowym, co często nie jest możliwe do zrealizowania w przypadku wykorzystania tradycyjnych macierzy blokowych i plikowych. Adresowanie wymagań na przykład dla takich regulacji, jak: RODO, Trwały Nośnik, MIFID2, SEC17a4f i wielu innych z zakresu prawa finansowego i podatkowego, służby zdrowia, bezpieczeństwa środowiskowego, ochrony danych oraz ich wyszukiwania i ujawniania (e-Discovery), ochrony konsumentów, itp.

”

Ważną funkcjonalnością Hitachi NAS, która pozwala na oszczędność przestrzeni dyskowej nawet do 90% jest wbudowana deduplikacja plików.”

Modernizacja analityki dla zbioru danych niestrukuralnych

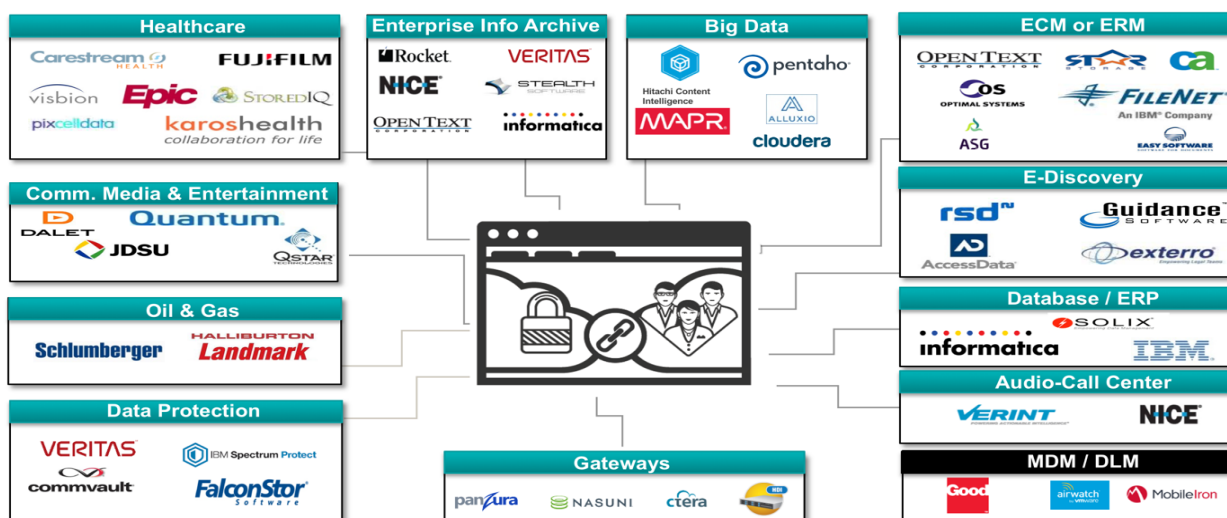
1. Wykorzystanie potencjału metadanych własnych, które są przechowywane wewnątrz macierzy obiektowej. Opisanie dokumentów za pomocą dodatkowych informacji oraz ich późniejsza analiza lub klasyfikacja w oparciu o ten metadane.
2. Wyszukiwanie i indeksowanie danych oraz metadanych. Ich klasyfikacja. Wykorzystanie wbudowanych mechanizmów oraz interfejsów programistycznych API dostępnych w macierzy obiektowej, do budowania środowiska dla operatorów i audytorów, którzy przeszukują i analizują zbiory danych niestrukuralnych. Integracja mechanizmów przeszukiwania z zewnętrznymi aplikacjami.

“

Ważną funkcjonalnością Hitachi NAS, która pozwala na oszczędność przestrzeni dyskowej nawet do 90% jest wbudowana deduplikacja plików.”

W powyższych scenariuszach obiektowy magazyn danych często jest jednym z kilku elementów całego rozwiązania. W architekturze korporacyjnej zastępuje on tradycyjną macierz plikową lub obiektową. Nie jest natomiast konkurentem dla aplikacji, w których operatorzy całego środowiska definiują kryteria i polityki dla określonych zbiorów danych które w tych aplikacjach mają być przechowywane.

Hitachi Content Platform posiada bardzo bogaty ekosystem aplikacji, z którymi potrafi współpracować i z którymi możliwa jest jego integracja realizowana albo za pomocą otwartych protokołów sieciowych, albo gotowych adapterów (pluginów). Dzięki takiej integracji w wielu przypadkach kryteria i polityki (np. retencji) definiowane przez operatora aplikacji dziedziczone są przez HCP i realizowane dodatkowo na poziomie sprzętowym.

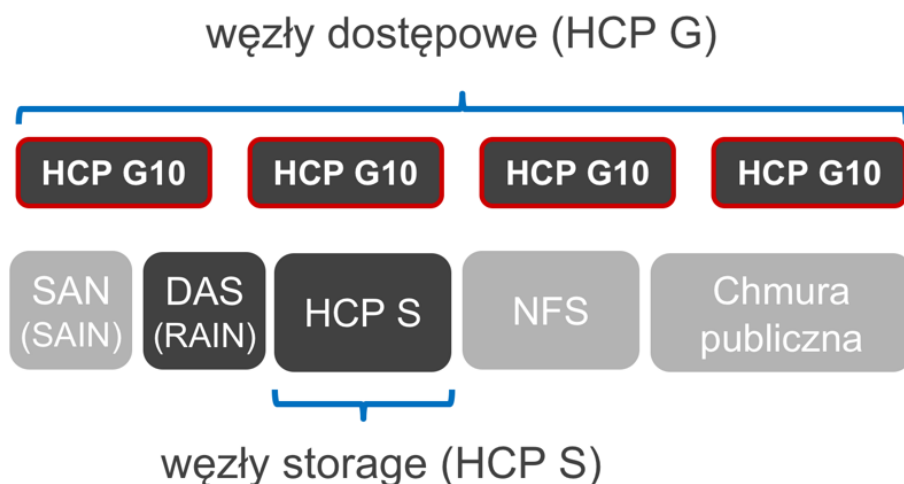


HCP i jego ekosystem aplikacji ISV.

Dodatkowo w **Hitachi Content Platform** dostępne są **interfejsy programistyczne API oraz biblioteki SDK**, które pozwalają na integrację w zasadzie z dowolną aplikacją, również taką, która jest przygotowywana i pisana na potrzeby konkretnego projektu i zastosowania. HCP posiada interfejsy API zarówno dla protokołów dostępowych, jak i do zarządzania (MAPI) oraz wyszukiwania w oparciu o wbudowany silnik Metadata Query Engine (MQE).

Budowa HCP

Hitachi Content Platform może być wdrażane albo w wirtualnych maszynach (HCP VM), albo na fizycznych serwerach dostępowych (HCP seria G). W obydwu tych przypadkach węzły klastra dostępowego (HCP VM lub HCP G) wirtualizują i centralizują przestrzeń dyskową dostarczaną wewnątrz każdego z tych serwerów, w węzłach typu storage HCP S, w macierzach dyskowych podłączonych do serwerów dostępowych, w serwerach plików podłączonych za pomocą protokołu NFS oraz z chmury publicznej.



Rysunek 10. Elastyczne możliwości dostarczania przestrzeni dyskowej w HCP.

**CHCESZ WIEDZIEĆ WIĘCEJ?
ZAPRASZAM DO KONTAKTU**

S4E S.A.

ul. Samuela Lindego 1 C,
30-148 Kraków
www.s4e.pl

HITACHI
Inspire the Next

