

Platforma aplikacyjna SAP HANA jako przykład komercyjnej realizacji technologii In-Memory Computing (IMC)

Marian Krupa, Maciej Pękala

Streszczenie

W niniejszym opracowaniu przedstawiono podstawową ideę przetwarzania w pamięci komputera zbiorów danych, ze szczególnym uwzględnieniem zastosowań biznesowych. Zaprezentowano rozwiązanie technologiczne SAP HANA wraz z krótkim opisem idei tej bazy danych w pamięci oraz dyskusją o innowacyjności. Zwrócono uwagę na efektywne przetwarzanie w czasie rzeczywistym bardzo dużych zbiorów danych. Przedyskutowano także znaczenie rozwiązania produktu SAP HANA dla biznesu oraz wskazano na korzyści, jakie można odnieść, w szczególność na przewagę konkurencyjną.

Wstęp

Wraz rozwojem technologii informatycznych w drugiej połowie XX wieku nastąpił gwałtowny wzrost zainteresowania przedsiębiorstw aplikacjami, które umożliwiały integrację i automatyzację procesów biznesowych. Rozwój zintegrowanych systemów informatycznych obejmujący fazy rozwiązań klasy MRP, ERP, SOA, BI ukształtował ostatecznie konieczność wypracowania sprawnego modelu zarządzania danymi przetwarzanymi w czasie rzeczywistym. Dostępne technologie umożliwiające przechowywanie i przetwarzanie danych na bazie tradycyjnych rozwiązań bazodanowych ograniczały możliwość pełnego wykorzystania najnowszych rozwiązań analitycznych klasy Business Intelligence.

Firma SAP, jako lider rozwiązań informatycznych dla biznesu, biorąc pod uwagę ograniczenie wynikające z klasycznych rozwiązań, podjęła się zadania wypracowania całkowicie nowego modelu technologicznego opartego na aplikacjach typu „in-memory”. Efektem tych prac, było zaprezentowanie w 2010 roku komercyjnej wersji aplikacji technologicznej o nazwie SAP HANA.

Celem jest opisanie technologii informatycznej In Memory Computing (IMC) zawartej w produkcie SAP HANA firmy SAP oraz wskazanie na klu-

czowe obszary związane z innowacyjnością i konkurencyjnością wynikające z jej wdrożenia w biznesie.

Zwrócono uwagę na rozwiązanie bazodanowe oparte na przechowywaniu dużej ilości danych w kolumnach. Realizacja tej idei stała się możliwa dzięki technologii IMC. Przedstawiono także różnicę pomiędzy strukturą bazy opartą o wiersze a strukturą opartą o kolumny. Istotnym efektem kolumnowej struktury bazy jest kilkukrotna kompresja danych, osiągnięta głównie poprzez usunięcie redundancji informacji.

Ewolucja technologii opartej na „In-Memory Computing (IMC)”

Rozwiązania typu „in-memory” sięgają już pierwszych zastosowań w zakresie budowania systemów informatycznych w oparciu o „Random Access Memory (RAM)”, tj. pamięci o dostępie swobodnym. „W tej pamięci zapisywany jest podczas pracy program, dane i niektóre niezbędne komputerowi do pracy informacje, które noszą nazwę zmiennych systemowych”¹. W przeciwieństwie do pamięci stałej „Read Only Memory (ROM)”, pamięć typu RAM (definiowana również jako pamięć operacyjna) umożliwia bardzo szybkie przetwarzanie danych, jednakże przy stosunkowo niewielkich ilościach danych i bez możliwości ich trwałego zapisu.

Pierwsze próby, prototypy wykorzystania technologii „in-memory” w różnych aplikacjach informatycznych pojawiły się już w końcu lat 90. ubiegłego wieku u różnych producentów, dostawców oprogramowania IT [Pezzini], [Gill]². Jednakże, dopiero w latach 2008-2010 rozpoczęto systematyczne prace nad całkowicie nowym, na wskroś innowacyjnym modelem architektury systemowej definiowanej jako „Im-Memoring Computing (IMC)”.

Współczesna definicja technologii informatycznej opartej na modelu „in-memory” nazywanej „In-Memory Computing (IMC)” jest formułowana w sposób następujący:

Metoda przetwarzania danych polegająca na masowym przeniesieniu danych podstawowych (*primary data*) z napędu dysku twardego (nośnik magnetyczny) do pamięci operacyjnej RAM [Pezzini]³.

¹ M. Pękala, *Wybrane zagadnienia informatyki*.

² J. Gill, *Shifting the BI Paradigm with In-Memory Database Technologies*.

³ M. Pezzini, *The Next Generation Architecture: In-Memory Computing*.

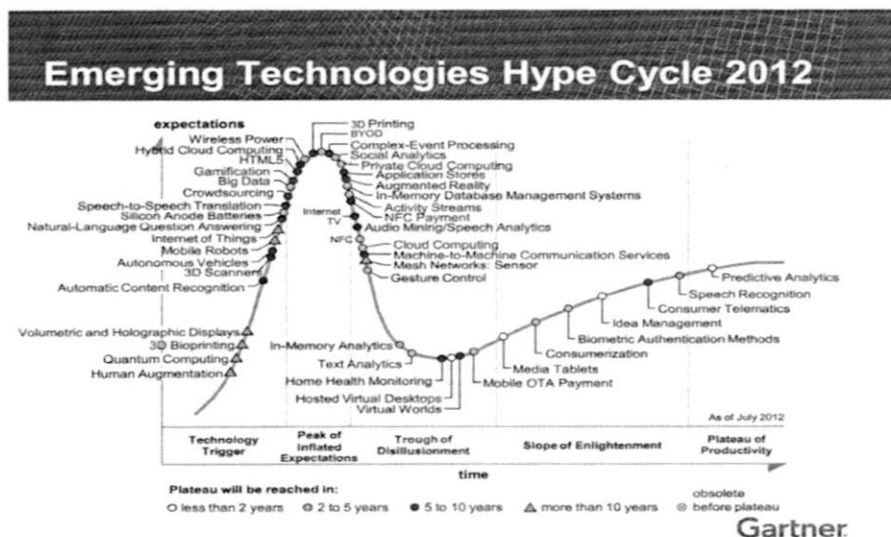
Przechowywanie informacji w pamięci operacyjnej RAM na dedykowanych do tego typu operacji serwerach w przeciwieństwie do wykorzystania skomplikowanych klasycznych rozwiązań bazodanowych wykorzystujących relatywnie powolne dyski twarde [Techopedia]⁴.

Wschodząca technologia (*emerging technology*), która zyskuje na powszechnej akceptacji, umożliwia użytkownikom natychmiastowy dostęp do właściwej informacji co bezpośrednio wpływa na jakość decyzji. Używając technologii „in-memory” dane są zaciągane do pamięci o dostępie swobodnym (RAM) zamiast do twardych dysków w związku z czym, pracownicy działów IT spędzają mniej czasu na prace deweloperskie zarówno w zakresie modelowania struktur danych, mechanizmów wyszukiwania danych, budowania kostek danych czy też projektowania tabel zawierających dane. [Earls]⁵.

W oparciu o przygotowywany cyklicznie raport firmy Gartner dotyczący analizy tzw. „*Emerging Technologies*”, technologia „In-Memory Computing” (IMC) w roku 2012 znajduje się w punkcie na styku fazy definiowanej jako „Peak of Inflated Expectations” oraz „Sliding Into the Trough”. Oznacza to, że IMC jest w wczesnej fazie rozwoju zarówno na płaszczyźnie technologicznej jak i komercyjnej. Jest to typowa sytuacja dla nowych i innowacyjnych technologii. Dlatego też, pierwsze wdrożenia nowych technologii, pierwsze projekty wdrożenia systemów informatycznych, w tym IMC charakteryzują zarówno spektakularne sukcesy jak też ryzyko potencjalnych porażek. Jednakże, w oparciu o wypracowany przez firmę Gartner model „Hype cycle for emerging technologies” (rys. 56) wraz z procesem udoskonalenia technologii jak i samej oferty biznesowej IMC w perspektywie 2–3 lat powinno wejść w fazę „Slope of Enlightenment”, tj. w fazę, w której to kolejne generacje rozwiązań, uzyskują powszechną akceptację ze strony biznesu. Przez co z fazy innowacji technologia przechodzi w fazę masowej implementacji.

⁴ „Techopedia”.

⁵ A. Earls, *Tips on evaluating, deploying and managing in-memory analytics tools*.



Rys. 56. Wschodzące technologie (*Emerging Technologies*) na tle tzw. „mega cyklu” (*Hype Cycle 2012*)⁶.

Obecnie rynek oprogramowania „in-memory”, zgodnie z oceną szacunkową firmy konsultingowej Gartner osiągnie w roku 2016 poziom 1 miliarda USD. Wynika to z konieczności zarówno: 1) wymiany przez organizacje biznesowe przestarzałej technologii służącej do przetwarzania dużej ilości danych i zastosowania bardziej konkurencyjnych i ekonomicznych rozwiązań opartych na technologii IMC jak też, 2) dostrzegania istotnych korzyści biznesowych wynikających z wdrożenia nowej technologii.

Temu procesowi sprzyjać będzie również stały spadek cen w zakresie wymaganej infrastruktury informatycznej, ze szczególnym uwzględnieniem pamięci operacyjnej typu DRAM i NAND. Zgodnie z opinią przedstawioną w cytowanym raporcie, technologia IMC stanowić będzie w najbliższych trzech latach, główny element strategii wzrostu wartości organizacji na całym świecie⁷.

⁶ *Gartner's Software Hype Cycles for 2012*, www.infoq.com/news/2012/08/Gartner-Hype-Cycle-2012 [18.07.2013].

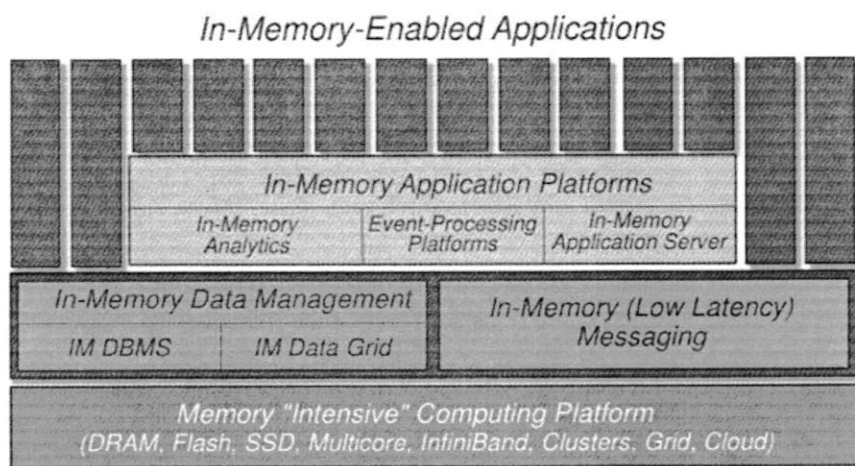
⁷ *Gartner Says In-Memory Computing Is Racing Towards Mainstream Adoption*, notatka opublikowana 3 kwietnia 2013, www.gartner.com/newsroom/id/2405315 [18.07.2013].

Systematyka rozwiązań informatycznych opartych o technologię „In-Memory Computing”

Fundamentem technologii IMC jest możliwość wykorzystania dużej ilości pamięci operacyjnej RAM (128 GiB⁸, 256 GiB, 512 GiB, 1 TiB lub więcej w zależności od wymaganej mocy serwera) w celach obliczeniowych, zbliżonych do czasu rzeczywistego (*Real-Time Computing*).

W ramach technologii IMC możemy wyróżnić następujące kategorie rozwiązań (rys. 57):

- programy informatyczne pracujące w oparciu o „in-memory”;
- platformy informatyczne realizujące zadania o charakterze analitycznym, transakcyjnym i sprzętowym (*Application Server*);
- zarządzania danymi (*DBMS, Data Grid*), jak też komunikatory w zakresie przesyłania danych w czasie rzeczywistym;
- platformy wspierające przetwarzanie olbrzymich ilości danych (*DRAM, Flash, SSD* itd.)

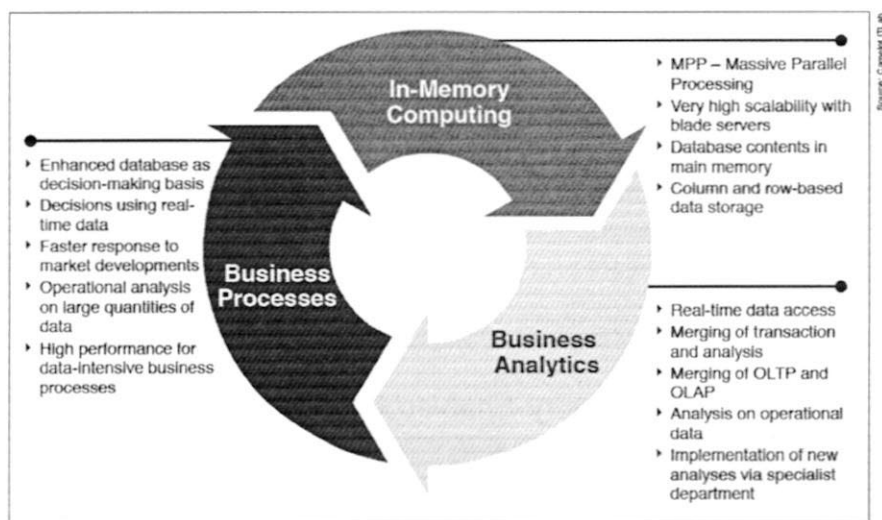


Rys. 57. Systematyka technologii opartych o model „In-Memory Computing” [Pezzini]⁹.

⁸ GiB, TiB oznaczenie mnożników binarnych ($Gi = 2^{30}$, $Ti = 2^{40}$) dla uniknięcia nieporozumienia przez użycie zapożyczonych z SI i stosowanych oznaczeń GB, TB co sugeruje mnożniki o podstawie 10 ($G = 10^9$, $T = 10^{12}$).

⁹ M. Pezzini, *op. cit.*

Do podstawowych elementów technologicznych i zarazem innowacyjnych rozwiązań informatycznych opartych na modelu IMC należy zaliczyć: „Multi-Core Memory Management (MCOMM)”, „Massive Parallel Processing (MPP)”, „Column Based Data Storage (CBDS)” oraz „Data Mart (DM)”. Wpisują się one w model wzajemnych integracji technologia (In-Memory Computing) – biznes (Business Analytics / Business Processes) (rys. 58).



Rys. 58. Model wzajemnych interakcji technologia – biznes [Merz, Simon]¹⁰.

Obecnie najbardziej znanymi rozwiązaniami opartymi na IMC są: SAP HANA (SAP), SolidDB (IBM), Times Ten (Oracle), Terracotta BigMemory (Software AG), MemSQL (MemSQL), Kognitio Analytical Platform (Kognitio), Manifact (Manifact). Należy jednak pamiętać, że nie wszystkie dostępne na rynku produkty spełniają w pełni standardy technologii IMC.

Do współczesnych rozwiązań sprzętowych pracujących w oparciu o technologię IMC należy zaliczyć m.in.: SAP HANA Appliance (SAP), Microsoft Parallel Data warehouse (Microsoft), Netezza Data Warehouse Appliance (IBM), Vertica Analytics Platform (HP), Active Enterprise Data Warehouse 5600 (Teradata), Oracle Exadata Database Machine (Oracle), Greenplum Data Computing Appliance (EMC).

¹⁰ M. Merz, T. Simon, *In-memory computing – a paradigm shift for the data warehouse*.

Innowacyjność technologii informatycznej opartej na nowoczesnych rozwiązaniach „In-Memory Computing”

Kluczowym aspektem innowacyjnym w technologii IMC jest tzw. „**Multi-Core Memory Management**” (MCMM). Zarządzanie pamięcią zarówno w oparciu o procesy wielordzeniowe jak i przetwarzanie danych za pomocą obliczeń równoległych pozwala na znaczny wzrost sprawności każdego systemu, ze szczególnym uwzględnieniem analiz prowadzonych na bardzo dużych zbiorach danych.

Niezwykle użytecznym elementem IMC jest również możliwość przechowywania dużej ilości danych w kolumnach (**Column Based Data Storage**), co umożliwia szybsze wyszukiwanie odpowiednich rekordów oraz efektywniejsze wykorzystanie mechanizmów kompresji danych.

Tabela 3. Przykład zastosowania systemu zarządzania danymi opartego na modelu „Column Based Data Storage”

Stock	Product	Sales
Germany	A	1,000
Germany	B	1,500
France	A	1,100

Row-oriented:

Row 1	Germany	A	1,000
Row 2	Germany	B	1,500
Row 3	France	A	1,100

Column-oriented:

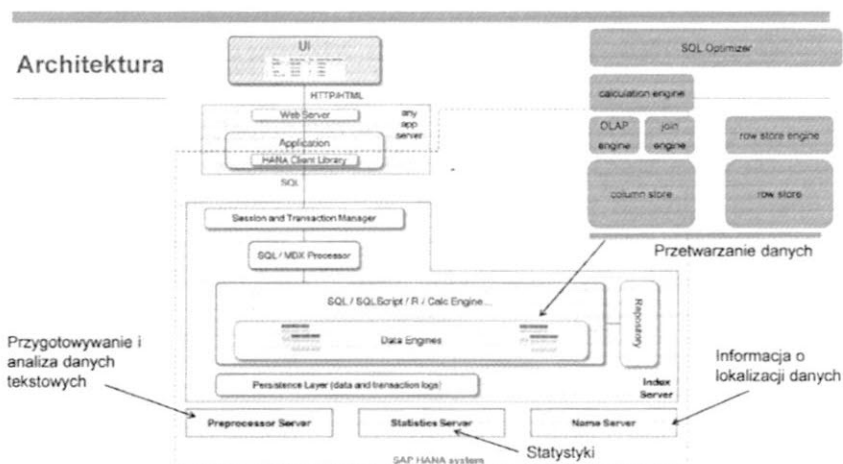
Column 1	Germany	Germany	France
Column 2	A	B	A
Column 3	1,000	1,500	1,100

Table: Difference between row-oriented and column-oriented data storage

Źródło: [Merz, Simon]¹¹.

W technologii SAP HANA zastosowano przetwarzanie kolumnowe, co pozwoliło na kilkukrotną kompresję danych oraz zastosowania technik analitycznych, które przy tradycyjnym dyskowym składowaniu danych wymagały wstępnej agregacji danych i stosowania technologii hurtowni danych. Schemat powyższego rozwiązania przedstawia rysunek 59.

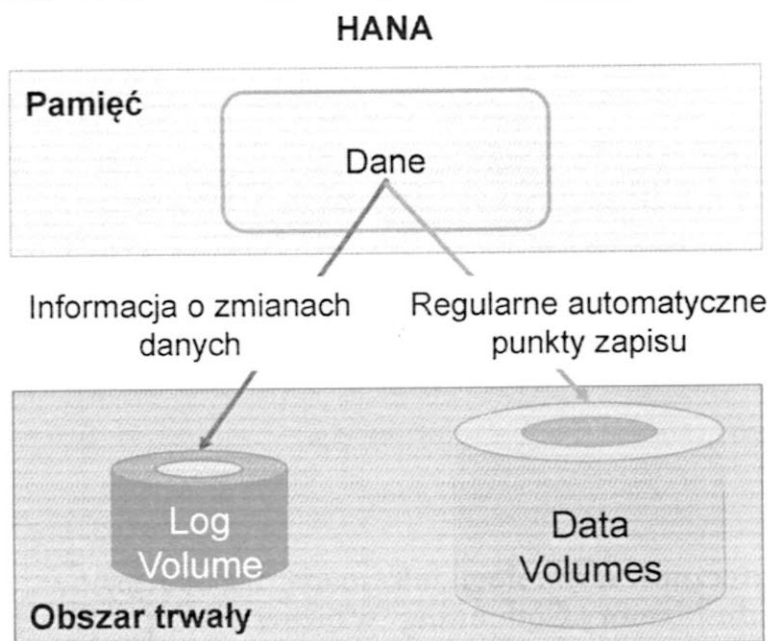
¹¹ M. Merz, T. Simon, *In-memory computing...*, op. cit.

Rys. 59. Architektura rozwiązania SAP HANA [Dworak]¹².

Warto zwrócić uwagę na fakt, że HANA jest kolumnową bazą danych rezydującą w pamięci, która realizuje wszystkie funkcje klasycznej bazy danych. Dane wprowadzane do systemu są wprowadzane wprost do pamięci i od razu mogą być wykorzystane zarówno w przetwarzaniu transakcyjnym jak i analitycznym. Ponieważ pamięć operacyjna (RAM) jest bardzo wrażliwa na zaniki pamięci powstaje oczywiste pytanie o bezpieczeństwo danych w bazie typu „in-memory”. W rozwiązaniu SAP HANA [Dworak]¹³ zastosowano szybkie dyski do bieżącego zapisu danych przechowywanych w pamięci. W pamięci przechowujemy dane oraz dodatkowe informacje takie jak np. modele danych, aby uzyskać maksymalną prędkość przetwarzania. Natomiast na dyskach stanowiących z pamięcią całość technologiczną dane zawarte w pamięci zapisywane są w dwóch obszarach: logów i punktów zapisu (ang. savepoint). Logi zawierają informacje o zmianach danych (redo log), są bezpośrednio zapisywane w sposób trwały po potwierdzeniu transakcji oraz są nadpisywane cyklicznie, czego warunkiem jest wykonanie backup'u. Zmienione dane oraz obszar „undo” zapisywane są do postaci trwałej w sposób cykliczny nie rzadziej niż co 5 minut, czas ten jest konfigurowalny. Rys. 5. przedstawia schemat techniczno-funkcjonalny rozwiązania.

¹² K. Dworak, *Architektura rozwiązania Business Suite na platformie SAP HANA*.

¹³ *Ibidem*.



Rys. 60. Bezpieczeństwo danych w rozwiązaniu SAP HANA [Dworak]¹⁴.

Kolejnym elementem w zakresie innowacji jest rozwiązanie typu „**Data Mart**” (tzw. „tematyczne hurtownie danych”). Powyższa funkcjonalność, w przeciwieństwie do typowych hurtowni danych, umożliwia grupowanie danych według kategorii obejmujących wybrane obszary działalności przedsiębiorstwa (oddziały, departamenty) takie jak: finanse, sprzedaż, logistyka, marketing itd. lub według uporządkowania procesowego czy też projektowego. Rozwiązania typu IMC umożliwiają budowanie tematycznych hurtowni danych zarówno opartych o klasyczne hurtownie danych lub całkowicie od nich niezależne. W pierwszym przypadku kluczową kwestią jest zapewnienie spójności pomiędzy hurtownią danych a „data mart”-ami w obszarze definicji, aktualizacji i zarządzania¹⁵.

W oparciu o dane branżowe możemy przedstawić dwa podejścia do przechowywania danych w „data mart”-ach¹⁶:

¹⁴ *Ibidem*.

¹⁵ *Data mart – tematyczna hurtownia danych.*

¹⁶ *Ibidem*.

„Data Mart”-y bazodanowe – lub też ekstrakty w postaci plików tekstowych – jednowymiarowe, niezagregowane zbiory danych. Bardzo często przetwarzanie i agregacja danych odbywa się po stronie aplikacji do raportowania.

Wielowymiarowe tematyczne hurtownie danych (MDDB) – dane sumaryczne zorganizowane w strukturze wielowymiarowej. Kalkulacje wykonywane są tylko raz w momencie tworzenia struktury wielowymiarowej i gotowe do analiz biznesowych w kategorii OLAP.

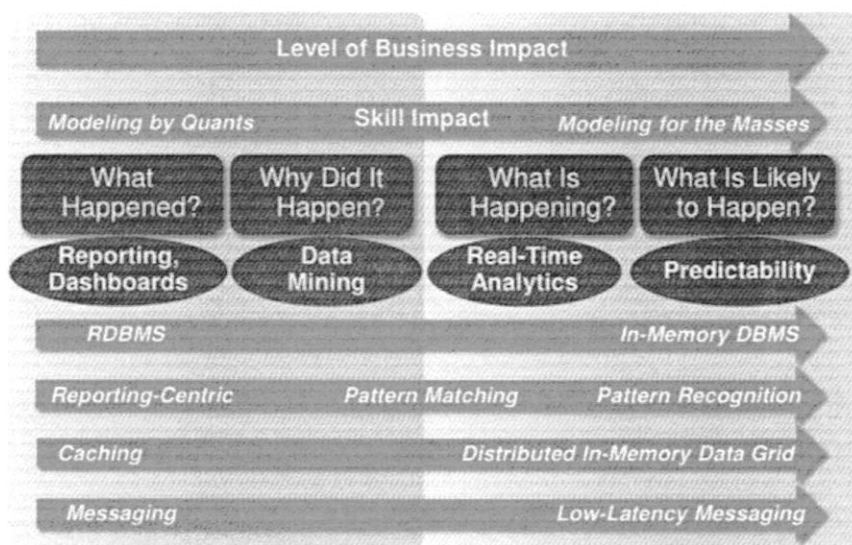
Pomimo wielu podobieństw w zakresie budowy architektury systemu, „data-mart” różni się w sposób zasadniczy od klasycznych hurtowni danych. Kluczowym aspektem jest metoda porządkowania danych, przez co te dwa systemy zarządzania danymi zawierają inną strukturę jak i sam zakres danych¹⁷.

Użyteczność technologii informatycznej opartej na „In-Memory Computing” w biznesie

Wdrożenie technologii informatycznej opartej na „In-Memory Computing” (IMC) umożliwia uzyskanie wielu korzyści biznesowych w zakresie opracowywania wielowymiarowych raportów i analiz w czasie rzeczywistym opartych przede wszystkim na następującej innowacji technologicznej: przetwarzanie wielkiej ilości danych następuje w pamięci operacyjnej bez konieczności wykorzystania w sposób tradycyjny dysków twardych.

Zakres innowacyjności technologii IMC jest bardzo szeroki. Z punktu widzenia potrzeb biznesu, na szczególną uwagę zasługuje zestaw funkcji analitycznych niedostępnych w klasycznych aplikacjach bazodanowych (rys. 61). Zawierają one możliwość już nie tylko zestawienia odpowiedzi na pytanie typu „Co się wydarzyło w firmie?” ale również znalezienie odpowiedzi typu: „Dlaczego się to wydarzyło?”, „Jak wygląda obecnie sytuacja?” czy też „Co prawdopodobnie wydarzy się w przyszłości?”. Na powyższe pytania rozwiązania analityczne oparte o IMC dostarczają odpowiedzi natychmiast, tzn. „w czasie rzeczywistym”. Wszelka modyfikacja samego zapytania oraz funkcja typu „drill-down” umożliwia uzyskanie odpowiedzi również w sposób natychmiastowy.

¹⁷ B. Inmon, *Data Mart Does Not Equal Data Warehouse*.



Rys. 61. Zestaw funkcji analitycznych niedostępnych w tradycyjnych aplikacjach baz danych [Pezzini]¹⁸.

W oparciu o przeprowadzone analizy (Gartner) wskazano, że w sposób szczególny zmiany technologiczne oparte o IMC będą dotyczyć takich sektorów gospodarki jak (rys. 62): reklama, bankowość, rynki kapitałowe, sektor energetyczny i usługi użyteczności publicznej, wywiad, transport i logistyka, media, wojskowość, gry on-line i rozrywka, sprzedaż, usługi typu SaaS (oprogramowanie jako usługa), usługi telekomunikacyjne oraz sprzedaż internetowa.

¹⁸ M. Pezzini, *op. cit.*

High	Advertising Banking Capital markets Energies and utilities Intelligence Transportation/logistics Media	Military Online gaming/entertainment Retail SaaS Telecom Web commerce
Medium	Automobile Central government CPG Healthcare Insurance Manufacturing Pharma	
Low	Construction Education Engineering Local government Mining	

Rys. 62. Wpływ IMC na poszczególne sektory gospodarki [Pezzini]¹⁹.

Na podstawie już zdobytych doświadczeń biznesowych oraz realizowanych prac projektowych można stwierdzić, że zakres zastosowania technologii IMC w biznesie jest praktycznie nieograniczony. Świadczą o tym kolejne aplikacje, które usprawniają pracę we wszystkich kluczowych obszarach funkcjonowania IT w organizacjach zarówno o strukturze korporacyjnej jak i sektora Małych i Średnich Przedsiębiorstw (MŚP).

SAP HANA (High Performance Analytic Appliance) – opis rozwiązania informatycznego firmy SAP opartego na technologii „In-Memory Computing”

Pierwsze prace obejmujące zastosowanie technologii „in-memory” w nowym wymiarze HANA (Hasso’s New Architecture) były prowadzone w sposób udokumentowany przez Hasso Plattner Institute i Stanford Uni-

¹⁹ *Ibidem*.

versity od 2008 roku²⁰. Próbowano opracować całkowicie nową architekturę oprogramowania, które umożliwiłoby przeprowadzania zaawansowanych i złożonych analiz danych w czasie rzeczywistym z wykorzystaniem technologii „in-memory”.

Producent, firma SAP, wykorzystał własne wcześniejsze doświadczenia w zakresie zastosowania technologii „in-memory” z takich pionierskich aplikacjach jak: *Business Warehouse Accelerator (BWA)*, *Advanced Planning and Optimization (APO)* czy też pakiet rozwiązań w ramach *CRM Segmentation, By Design (analytics)* i *Enterprise Search* oraz rodziny aplikacji SAP Sybase.

Ostatecznie, całkowicie nowa platforma SAP HANA (High Performance Analytic Appliance²¹), stanowiąca obecnie podstawę oprogramowania firmy SAP, została po raz pierwszy zaprezentowana publicznie 20 listopada 2010 roku jako produkt komercyjny w wersji SP0 w trakcie SAP TechEd 2010 w Berlinie. Jest to data, która stanowi punkt odniesienia do kolejnych wersji oznaczanych w kategoriach od SP0 do SP6 – tj. obecny poziom tzw. „Support Package Release” opublikowany w dniu 28 czerwca 2013 roku²². Pierwsze komercyjne wdrożenie SAP HANA nastąpiło już w grudniu 2011 w ramach projektu „Strategic Workforce Planning”²³.

Należy zwrócić szczególną uwagę na fakt, zgodnie z publiczną wypowiedzią Hasso Plattnera (współzałożyciela oraz przewodniczącego rady nadzorczej firmy SAP – stan na rok 2012), że rozwiązanie SAP HANA stanowi „de facto” tzw. „versionless” system”, czyli oprogramowanie, który podlega jedynie systematycznie planowanym uaktualnieniom, a nie klasycznemu wersjonowaniu uzyskanemu poprzez zastosowanie uciążliwych działań w postaci wdrażania tzw. „upgradów”. Do dnia 28 czerwca 2013 roku firma SAP opublikowała następujące „support packages” (uaktualnienia) dla platformy SAP HANA²⁴:

- SP0 – 20 listopada 2010; pierwsza publiczna prezentacja SAP HANA;
- SP1 – 20 czerwca 2011; ogólnie dostępny produkt komercyjny SAP HANA (tzw. „General Availability” – GA);
- SP2 – 27 czerwca 2011; rozwój funkcji tzw. „Data Mart” (tematyczna hurtownia danych);

²⁰ S. Vishal, *Timeless Software*.

²¹ Desmond, *SAP HANA – Updating the Naming Conventions*.

²² *Unveiling SAP TechEd 2010*.

²³ <http://timoelliott.com/blog/2010/12/sap-introduces-first-hana-enabled-application-links-and-screen-shots.html>.

²⁴ J. Appleby, *Updated: The SAP HANA FAQ - answering key SAP In-Memory questions*.

- SP3 a.k.a HANA 1.5 – 7 listopada 2011; rozszerzenie HANA dotyczące integracji z hurtownią danych SAP Business Warehouse (BW); tzw. „Project Orange”;
- SP4 – maj 2012; kolejne rozszerzenie o nowe funkcje w ramach SAP BW;
- SP5 – listopad 2012; prezentacja “Extended Application Services (REST driver)”;
- SP6 – 28 czerwca 2013; prezentacja najnowszego „support packages” dla SAP HANA.

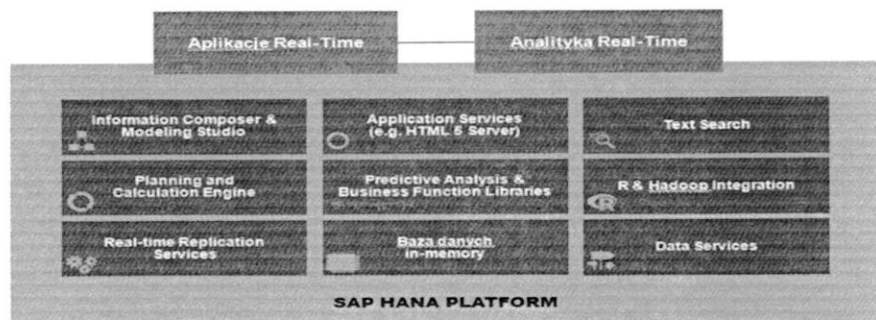
W ramach poszczególnych „support packages” i niezależnie od rozwoju samego rdzenia platformy SAP HANA (czy też HANA DB - rozwiązanie obejmujące pracę w ramach hurtowni danych) firma SAP dodatkowo opracowała w latach 2011–2013 m.in. takie komponenty jak: SAP HANA Studio – zestaw narzędzi służących do konfiguracji i modelowania danych; SAP HANA Appliance – serwery certyfikowane przez firmę SAP, dedykowane do obsługi rozwiązań SAP HANA DB, jak też zestaw narzędzi do transformacji i replikacji danych; SAP HANA One – dostosowanie platformy HANA do jej wykorzystania w ramach Amazon Web Services (AWS) w wersji SAP HANA Enterprise Cloud – infrastruktura umożliwiająca pracę w technologii tzw. „chmury” (Cloud Computing) dostosowana do platformy HANA.

W Polsce pierwsza umowa na wdrożenie platformy SAP HANA została podpisana 5 kwietnia 2011 roku przez firmę SAP Polska oraz Grupę CEDC - wiodącego światowego producenta wyrobów alkoholowych, notowany na giełdach w Nowym Jorku i Warszawie. Zgodnie z opublikowaną notatką prasową „CEDC zamierza wykorzystać rozwiązanie SAP HANA w tworzeniu spójnej platformy hurtowni danych do celów przetwarzania danych finansowych i raportowania. Wykorzystanie rozwiązania in-memory firmy SAP ma na celu zwiększenie wydajności i szybkości raportowania, a także uproszczenie procesów konsolidacji finansowej”²⁵.

SAP HANA jest nowoczesną platformą informatyczną do masowego przetwarzania danych w czasie rzeczywistym opartą o technologię In-Memory Computing (IMC).

Wpisują się one w „otwartą” architekturę HANA, która umożliwia integrację z wieloma aplikacjami zarówno firmy SAP jak też „non-SAP” (rys. 8).

²⁵ Grupa CEDC wdroży rozwiązanie „in-memory” SAP HANA.



Rys. 63. Platforma SAP HANA²⁶.

Platforma HANA może być dostarczana zarówno w oparciu o stacjonarne serwery w ramach SAP HANA Appliance jak też w technologii „chmury” – SAP HANA Enterprise Cloud.

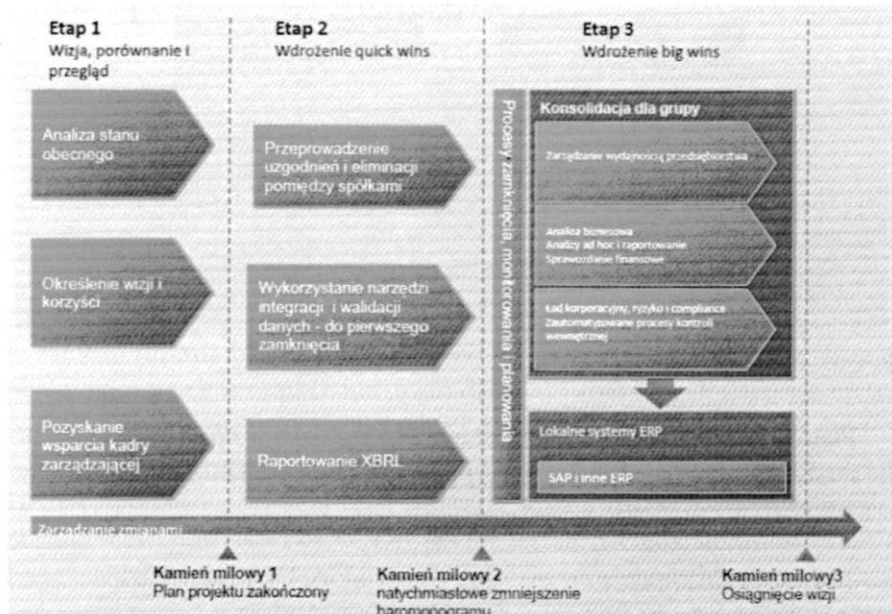
Metodyka wdrożenia rozwiązania SAP HANA – studium przypadku projektu wdrożenia SAP HANA w Grupie CEDC

Realizacja wdrożenia każdego systemu informatycznego, w tym SAP HANA wymaga odpowiedniego zaplanowania i potem realizacji. Profesjonalizm w powyższym zakresie jest kluczowym elementem zapewniającym uzyskanie sukcesu we wdrożeniu nie tylko w zakresie celów projektowych, informatycznych ale przede wszystkim biznesowych.

W oparciu o doświadczenie wypracowane przez firmę SID Group²⁷ w trakcie realizacji projektu wdrożenia SAP HANA w Grupie CEDC możemy wskazać na trzy etapy prac wdrożeniowych. Każdy z tych etapów kończy się precyzyjnie zaplanowanym i zdefiniowanym efektem, określanym w praktyce projektowej jako „kamień milowy” (rys. 64).

²⁶ SAP HANA oraz SAP HANA Platform – Technical Overview Driving Innovations in IT and in Business with In-Memory Computing Technology.

²⁷ <http://blog.sidgroup.pl/sid-group-jako-pierwsza-firma-w-polsce-wdraza-rozwiazanie-sap-hana>.



Rys. 64. Plan wdrożenia rozwiązania SAP HANA w Grupie CEDC²⁸.

Pierwszy etap zawiera opracowanie wizji biznesowej dla zadań informatycznych, co umożliwi precyzyjną ocenę sukcesu lub porażki realizowanego przedsięwzięcia. Do podstawowych zadań w tym etapie jest opracowanie wizji zmian oraz oczekiwanych korzyści biznesowych oraz pozyskanie wsparcia kadry zarządzającej, tzw. sponsora, w zakresie sprawnego zarządzania projektem – aspekt decyzyjny. Etap pierwszy kończy się podpisanym dokumentem definiowanym jako koncepcja biznesowa projektu wdrożenia nowego rozwiązania.

Etap drugi obejmuje realizację działań wdrożeniowych w tych obszarach, które zapewniają szybkie efekty (tzw. „quick wins”) w postaci wymierzonych korzyści biznesowych. W przypadku wdrożenia SAP HANA w Grupie CEDC wskazano na trzy aspekty: 1) przeprowadzenie uzgodnień i eliminacji korespondencji między spółkami; 2) wykorzystanie narzędzi integracji i walidacji danych oraz 3) raportowanie w formacie XBRL²⁹. Etap drugi kończy się zamknięciem części zadań wpisanych w harmonogram wdrożenia.

²⁸ P. Majewski, J. Bugajski, M. Sawczuk, *SAP HANA. Nowoczesna analityka i wsparcie dla konsolidacji finansowej*.

²⁹ eXtensible Business Reporting Language.

Etap trzeci wskazuje na działania mające na celu realizację kluczowych założeń (wizji) projektu (tzw. „big wins”). W tym etapie są realizowane zadania, które generują największą wartość ale równocześnie są najbardziej złożone i wymagają pełnego zaangażowania zarówno ze strony wykonawcy jak i zamawiającego. W przypadku projektu wdrożenia w Grupie CEDC głównym celem było opracowanie systemu konsolidacji całej grupy ze szczególnym zwróceniem uwagi na aspekty raportowania i analiz biznesowych. Efektem tego etapu, obok formalnych kwestii odbioru poszczególnych elementów nowego systemu było osiągnięcie wizji zapisanej w etapie pierwszym w postaci koncepcji biznesowej.

Analiza innowacyjności i konkurencyjności pakietu SAP HANA

Zdobyte już doświadczenie w zakresie wdrożeń SAP HANA opartych o technologię In-Memory Computing (IMC) wskazuje na dwa typy innowacji, tj. **innowacje sprzętowe** i innowacje obejmujące **oprogramowanie SAP**.

W pierwszym przypadku – innowacje sprzętowe (SAP HANA Appliance) – dają możliwość wykorzystania przede wszystkim wielordzeniowej architektury (multi-core database), co umożliwia bardzo dużą przepustowość danych na poziomie 100 GB/s przy równoczesnym dramatycznym spadku ceny w stosunku do wydajności, co jest niezwykle istotne z punktu widzenia planowanych inwestycji w obszary IT³⁰.

W zakresie innowacji w obszarze samego oprogramowania uzyskujemy takie funkcje jak: równoczesne zastosowanie kolumnowych i wierszowych struktur baz danych, znacząca poprawa kompresji danych, lepsze partycjonowanie, brak tabel zagregowanych jak też ładowanie danych przez zmianę (przyrost), tzw. „deltę”³¹.

Do kluczowych zalet wynikających z wdrożenia SAP HANA w organizacji możemy zaliczyć^{32 33 34}:

- umożliwia uzyskanie na jednej platformie analityki, danych operacyjnych oraz zarządzania strategicznego;

³⁰ *Ibidem*.

³¹ *Ibidem*.

³² *Ibidem*.

³³ M. Salwin, *Kilka słów na temat SAP HANA*.

³⁴ *SAP Product Road Map SAP HANA / Road Map Revision*.

- jest rozszerzeniem obecnych systemów, a nie zastępuje ich;
- umożliwia podejmowanie lepszych decyzji biznesowych poprzez zwiększenie widoczności dużej ilości danych biznesowych;
- o wiele szybsza reakcja na zdarzenia biznesowe – krótszy czas dzięki analizie w czasie rzeczywistym i raportowaniu danych operacyjnych,
- większa elastyczność, dzięki innowacyjnym scenariuszom analityki i raportowania w czasie rzeczywistym;
- znaczne skrócenie czasów potrzebnych na wykonywanie raportów w systemach klasy BI, w których zastosowano SAP HANA;
- możliwość zastosowania jako samodzielnego elementu pełniącego rolę hurtowni danych;
- brak konieczności zakupu licencji oprogramowania serwera baz danych w porównaniu do architektury wykorzystującej BW jako hurtownię danych;
- łatwa integracja z innymi systemami SAP.

Ponadto, w oparciu o przeprowadzone badania firmy SAP opracowano zestawienie dziesięciu powodów, dla których klienci wybierają właśnie rozwiązanie SAP HANA³⁵:

- **Szybkość** – możliwość zarządzania dużymi wolumenami danych z niewiarygodną szybkością;
- **elastyczność** – możliwość korzystania z informacji w czasie rzeczywistym, co bezpośrednio umożliwia uzyskanie przewagi konkurencyjnej;
- **dowolne źródło danych** – możliwość uzyskania informacji z danych strukturalnych i niestukturalnych;
- **wgląd** – wykorzystanie nowych możliwości pozyskiwania informacji na podstawie złożonej analizy predykcyjnej;
- **aplikacje nowej generacji** – uzyskanie przewagi technologicznej;
- **chmura** – dostęp do jednej z najbardziej zaawansowanych platform przetwarzania danych w tzw. „chmurze”;
- **innowacyjność** – zastosowanie optymalnej platformy dla inicjowania innowacji w biznesie;
- **prostota i ekonomiczność rozwiązania** – zarządzanie mniejszą liczbą warstw i prostszą infrastrukturą, obniża się dzięki temu koszty;
- **wartość** – wprowadzanie nowych rozwiązań bez zakłóceń działalności przy równoczesnym zwiększaniu wartości wcześniejszych inwestycji w IT;
- **wyбір** – możliwość pracy z wybranymi partnerami wdrożeniowymi w danym obszarze.

³⁵ Top 10. Dziesięć powodów, dla których klienci wybierają rozwiązanie SAP HANA.

W zakresie zarządzania dużymi wolumenami danych z niewiarygodną szybkością przeprowadzone badania wskazują wzrost na poziomie 3600 razy szybciej niż przed wdrożeniem SAP HANA (rys. 65).



Rys. 65. Zestawienie wyników badania w zakresie wykorzystania platformy SAP HANA³⁶.

Wyniki klientów mogą być inne i zależeć od konfiguracji systemu. Na rysunku 8, po lewej: Na podstawie wyników testów przeprowadzonych przez SAP z użyciem danych klienta z branży budowlanej pochodzących z transakcji w oprogramowaniu SAP Business Suite. Zapytania były przetwarzane średnio 3600 razy szybciej niż w oprogramowaniu SAP Business Suite z tradycyjnymi, dyskowymi bazami danych. Wyniki te osiągnięto przy zastosowaniu oprogramowania SAP HANA 1.0 na jednym czteroprocessorowym serwerze z 8 rdzeniami (częstotliwość taktowania 2,27 GHz), z 0,5 TB pamięci głównej, dyskiem SSD o pojemności 2 TB oraz kartą sieci Ethernet o szybkości transmisji 1 Gb/s, z systemem operacyjnym typu open source. Na rysunku 8, po prawej: Na podstawie wyników testów przeprowadzonych przez SAP z użyciem 460 miliardów wierszy danych z punktów sprzedaży, dostarczonych przez klienta będącego czołową dostawcą artykułów konsumpcyjnych. Wyniki te osiągnięto, przetwarzając typowe zapytania biznesowe na 10-węzłowej platformie z oprogramowaniem SAP HANA 1.0. Każdy węzeł stanowił jeden czteroprocessorowy serwer z 8 rdzeniami (częstotliwość taktowania 2,27 GHz), z 0,5 TB pamięci głównej, dyskiem SSD o pojemności 2 TB oraz kartą sieci Ethernet o szybkości transmisji 10 Gb/s, z systemem operacyjnym typu open Source.

³⁶ *Ibidem*.

³⁷

To co wyróżnia platformę HANA w porównaniu z klasycznymi rozwiązaniami bazodanowymi to: 1) możliwość integracji danych zarówno OLTP jak OLAP³⁸; 2) kolumnowa agregacja danych z wysokim poziomem kompresji; 3) 100% przetwarzanie danych w pamięci DRAM; 4) uproszczona liniowa struktura danych (*Stack*) (rys. 66).



Rys. 66. Porównanie platformy SAP HANA do klasycznych rozwiązań bazodanowych³⁹.

Kolejny przykład korzyści z zastosowania SAP HANA znajdujemy w referacie Anny Różyckiej⁴⁰, w którym m.in. przedstawiono wyniki testu firmy SAP. Test prowadzony był w oparciu o harmonogramowanie 763 000 planów prac w obszarze budownictwa. W klasycznej technologii z bazą zapisaną na dyskach pierwsza próba skończyła się wynikiem „time out”, system nie był w stanie jej zrealizować. Podzielono zatem prace na grupy po 100 000, co system potrafił już przetworzyć. Sumaryczny czas realizacji zadania wyniósł 17 godzin. Na platformie HANA pierwszy przebieg dla wszystkich 763 000 planów trwał 30 minut, natomiast kolejne już tylko 48–50 sekund!

W wymiarze korzyści biznesowych SAP HANA wspiera organizacje m.in. w następujących obszarach⁴¹:

Bardziej efektywne zarządzanie energią – przedsiębiorstwa użyteczności publicznej wykorzystują rozwiązanie SAP HANA do przetwarzania i analizy

³⁸ OLTP – On-line Transaction Processing, OLAP – On-line Analytical Processing.

³⁹ *Top 10 – Why SAP HANA.*

⁴⁰ A. Różycka, *SAP HANA dla biznesu. Czy to się może opłacać.*

⁴¹ *Faster, Higher, Stronger: In-Memory Computing Description and What SAP Means for Your Organization.*

dużych ilości danych generowanych przez inteligentne mierniki, poprawy efektywności energetycznej i wdrażania inicjatyw zrównoważonego rozwoju.

Zarządzanie trasami przejazdu w planowaniu sieci logistycznych w wymiarze transportu i spedycji – SAP HANA pomaga firmom w obliczaniu w czasie rzeczywistym optymalnych tras przejazdu przy użyciu danych GPS przesyłanych z tysięcy pojazdów.

Wykrywanie nielegalnego oprogramowania i wdrażanie środków zapobiegawczych – Firmy przemysłu technologicznego używają rozwiązania SAP HANA do analizy dużych ilości złożonych danych dotyczących nielegalnego oprogramowania, opracowywania środków zapobiegawczych i odzyskiwania utraconych przychodów⁴².

Przykłady biznesowych korzyści SAP HANA możemy również odnaleźć w wdrożeniach u ponad 1300 klientów na całym świecie⁴³. Są to korzyści, definiowane jako „business value”, które dotyczą przede wszystkim pełnej przejrzystości i kontroli finansowej dostępnej w czasie rzeczywistym, skrócenie czasu generowania złożonych raportów i analiz, bardziej inteligentne procesy decyzyjne, znaczne przyspieszenie wszystkich aplikacji biznesowych, w tym SAP ERP, lepsze zrozumienie potrzeb klienta jak też wysoki prestiż i uznanie w zakresie wdrażania innowacyjnych technologii⁴⁴.

Podsumowanie

Rozwiązanie informatyczne SAP HANA oparte na technologii definiowanej jako „in-memory” czy też „In-Memory Computing (IMC)” umożliwia podmiotom gospodarczym skuteczne zwiększenie wartości, jakości, konkurencyjności swoich produktów i usług, jak też tworzenie nowych lub udoskonalonych produktów i usług informatycznych, ze szczególnym uwzględnieniem wielowymiarowych analiz biznesowych tworzonych w czasie rzeczywistym, na bazie m.in. zintegrowanych aplikacji informatycznych klasy SAP BI (Business Intelligence), SAP NetWeaver – Business Warehouse działających na fundamencie technologicznym SAP NetWeaver. Dodatkową wartością biznesową jest również możliwość bardziej efektywnego zarządzania energią, infrastrukturą logistyczną oraz skutecznego monitorowania stosowania nielegalnego oprogramowania i wsparcia wdrażania środków zapobiegawczych w tym zakresie. SAP HANA w wer-

⁴² Top 10. Dziesięć powodów..., *op. cit.*

⁴³ SAP HANA, *op. cit.*

⁴⁴ Przykłady wdrożeń SAP HANA oraz referencje są dostępne na: <http://www54.sap.com/pc/tech/in-memory-computing-hana/customer-reviews.html>.

sji SAP HANA Enterprise Cloud” umożliwia pracę również w technologii tzw. „chmury” (Cloud Computing).

Reasumując, rozwiązanie informatyczne SAP HANA oparte na technologii definiowanej jako „in-memory” czy też „In-Memory Computing (IMC)” umożliwia podmiotom gospodarczym skuteczne zwiększenie wartości, jakości, konkurencyjności swoich produktów i usług, jak też tworzenie nowych lub udoskonalonych produktów i usług informatycznych, ze szczególnym uwzględnieniem wielowymiarowych analiz biznesowych tworzonych w czasie rzeczywistym.

Literatura

1. Pękala M., *Wybrane zagadnienia informatyki*, KSW, Kraków 2006, s. 19.
2. Gill J., *Shifting the BI Paradigm with In-Memory Database Technologies*, „Business Intelligence Journal”, 1 kwietnia 2007, pobrano z High Beam Research (<http://www.highbeam.com>), w dniu 23.07.2013.
3. Pezzini M., *The Next Generation Architecture: In-Memory Computing*, prezentacja firmy Gartner dostępna na: http://www.slideshare.net/SAP_Nederland/the-next-generation-architecture-inmemory-computing-massimo-pezzini w dniu 10.07.2013.
4. “Techopedia”, pobrana z <http://www.techopedia.com/definition/28539/in-memory-computing> w dniu 23.07.2013.
5. Earls A., *Tips on evaluating, deploying and managing in-memory analytics tools*, zamieszczono 11.10.2011, pobrano z <http://docs.media.bitpipe.com> w dniu 23.07.2013.
6. *Gartner’s Software Hype Cycles for 2012*, www.infoq.com/news/2012/08/Gartner-Hype-Cycle-2012 z dnia 18.07.2013.
7. *Gartner Says In-Memory Computing Is Racing Towards Mainstream Adoption*, notatka opublikowana 3 kwietnia 2013, www.gartner.com/newsroom/id/2405315 z dnia 18.07.2013.
8. Merz M., Simon T., *In-memory computing – a paradigm shift for the data warehouse?*, przedruk z www.sap-port.de z dnia 03.02.2012, pobrano z <http://www.camelot-itlab.com> w dniu 23.07.2013.

9. Dworak K., *Architektura rozwiązania Business Suite na platformie SAP HANA*, Konferencja SAP Business Suite na platformie HANA, Warszawa, 22.10.2013.
10. *Data mart – tematyczna hurtownia danych*, opracowanie branżowe dostępne na stronie: <http://etl-tools.info/pl/data-mart.html> w dniu 18.07.2013.
11. Inmon B., *Data Mart Does Not Equal Data Warehouse*, opracowanie branżowe dostępne na stronie: <http://www.information-management.com/infodirect/19991120/1675-1.html>, w dniu: 18.07.2013.
12. Vishal S., *Timeless Software*. „Timelessness” tzw. „personal blog” prowadzony przez współautora rozwiązania SAP HANA, informacja opublikowana 19 stycznia 2012 na stronie <http://vishalsikka.blogspot.com/2008/10/timeless-software.html>.
13. Desmond P., *SAP HANA – Updating the Naming Conventions*, [w:] „ERP Executive”, Sierpień 2011, www.erpexecutive.com.
14. *Unveiling SAP TechEd 2010*, materiały branżowe firmy BayForce (certyfikowanego partnera firmy SAP, www.bayforce.com/2010/10/unveiling-sap-teched-2010/ z dnia 17.07.2013.
15. <http://timoelliott.com/blog/2010/12/sap-introduces-first-hana-enabled-application-links-and-screen-shots.html>.
16. Appleby J., *Updated: The SAP HANA FAQ – answering key SAP In-Memory questions*, informacja opublikowana 23 stycznia 2012 na stronie www.bluefinsolutions.com.
17. *Grupa CEDC wdroży rozwiązanie „in-memory” SAP HANA*, notatka prasowa zamieszczona na stronie internetowej firmy SAP, www.sap.com/poland/press.epx?pressid=15060 z dnia 16.07.2013.
18. *SAP HANA*, SAP Polska, Warszawa 2013.
19. *SAP HANA Platform – Technical Overview Driving Innovations in IT and in Business with In-Memory Computing Technology*, materiały branżowe firmy SAP, Walldorf 2012.
20. <http://blog.sidgroup.pl/sid-group-jako-pierwsza-firma-w-polsce-wdraza-rozwiazanie-sap-hana>.
21. Majewski P., Bugajski J., Sawczuk M., *SAP HANA. Nowoczesna analityka i wsparcie dla konsolidacji finansowej*, prezentacja przedstawiona w trakcie SAP World Tour 2011, Sopot 9–10 czerwca 2011.

22. Salwin M., *Kilka słów na temat SAP HANA*, „Blog technologiczny ATENA”, <http://blog.atena.pl/kilka-slow-na-temat-sap-hana> z dnia 15.07.2013.
23. *SAP Product Road Map SAP HANA / Road Map Revision*, materiały branżowe firmy SAP z dnia 13.12.2012.
24. *Top 10. Dziesięć powodów, dla których klienci wybierają rozwiązanie SAP HANA*, materiały branżowe firmy SAP, 2013.
25. *Top 10 – Why SAP HANA*, materiały branżowe firmy SAP, lipiec 2013.
26. Różycka A., *SAP HANA dla biznesu. Czy to się może opłacać?*, Konferencja SAP Business Suite na platformie HANA, Warszawa, 22.10.2013.
27. *Faster, Higher, Stronger: In-Memory Computing Description and What SAP Means for Your Organization*, Opracowanie techniczne IDC sponsorowane przez firmę SAP AG, listopad 2011.