**Proof of Concept (PoC): Architektura Big Data dla Rockstar Games**

**Cel:**

Przetwarzanie danych w celu wykrycia i zapobiegania błędom w grze oraz zwiększenia satysfakcji użytkowników.

**Architektura Big Data:**

1. **Zbieranie danych:**
   * **Narzędzia:**
     + Azure Blob Storage: Do przechowywania danych gier, takich jak logi serwerów, dane graczy, zdarzenia w grze.
     + Azure Event Hubs: Do gromadzenia strumieniowych danych w czasie rzeczywistym, takich jak zdarzenia z gry, interakcje użytkowników.
   * **Compute:**
     + Azure Virtual Machines: Wykorzystane do przetwarzania danych zebranych w czasie rzeczywistym, przetwarzania wstępnego itp.
2. **Przetwarzanie danych:**
   * **Narzędzia:**
     + Azure Stream Analytics: Do analizy strumieniowej danych w czasie rzeczywistym w celu wykrywania anomalii i błędów w grze.
     + Azure Databricks: Do przetwarzania i analizy dużych zbiorów danych offline, jak również do bardziej zaawansowanych analiz predykcyjnych.
   * **Compute:**
     + Azure Databricks: Umożliwia skalowalne przetwarzanie danych oraz wykonywanie złożonych obliczeń.
3. **Przechowywanie danych:**
   * **Narzędzia:**
     + Azure Cosmos DB: Do przechowywania danych w czasie rzeczywistym, zapewniająca wysoką dostępność i niskie opóźnienia.
     + Azure SQL Database: Do przechowywania danych historycznych, statystyk, wyników gier.
   * **Compute:**
     + Azure Cosmos DB: Zapewnia natywną obsługę globalnie rozproszonych danych i skalowalność horyzontalną.
     + Azure SQL Database: Umożliwia przetwarzanie zapytań SQL oraz analizę danych z wykorzystaniem procedur składowanych i funkcji.
4. **Analiza danych i uczenie maszynowe:**
   * **Narzędzia:**
     + Azure Machine Learning: Do tworzenia i trenowania modeli uczenia maszynowego do przewidywania zachowań graczy, wykrywania anomalii w grze.
   * **Compute:**
     + Azure Machine Learning Compute: Zapewnia skalowalne zasoby obliczeniowe do treningu modeli uczenia maszynowego.
5. **Wizualizacja danych:**
   * **Narzędzia:**
     + Power BI: Do tworzenia interaktywnych raportów i wizualizacji danych, pozwalających na analizę statystyk gier, trendów, zachowań użytkowników.
   * **Compute:**
     + Azure Virtual Machines: Umożliwia wykonanie obliczeń niezbędnych do generowania raportów i wizualizacji.

**Wykorzystanie:**

1. Dane gromadzone z serwerów gier oraz interakcji użytkowników są strumieniowane do Azure Event Hubs.
2. Azure Stream Analytics przetwarza te dane w czasie rzeczywistym, identyfikując anomalie i błędy.
3. Dane są archiwizowane w Azure Blob Storage i Azure Cosmos DB dla dalszej analizy.
4. Azure Databricks wykorzystywane jest do przetwarzania offline danych oraz bardziej zaawansowanych analiz.
5. Azure Machine Learning służy do trenowania modeli uczenia maszynowego, które wykrywają anomalie i przewidują zachowania graczy.
6. Power BI wykorzystywane jest do wizualizacji danych i generowania raportów, umożliwiających zarządzanie grą i poprawę satysfakcji użytkowników.