



343 Industries Gets New User Insights from Big Data in the Cloud

Overview

Country or Region: United States
Industry: Professional services – IT services

Customer Profile

The 343 Industries development team hosts and manages Halo 4, an award-winning game for the Microsoft Xbox 360 video game and entertainment console.

Business Situation

The team needed to provide business intelligence (BI) insight about the game to internal and external customers.

Solution

The team implemented a solution that uses Windows Azure HDInsight Service, based on the Apache Hadoop data-processing framework, and Microsoft BI technologies.

Benefits

- Enhances user experience through increased agility and faster response times.
- Connects Halo 4 team directly to customers through weekly updates.
- Keeps playing field level by providing in-game analysis to detect cheaters.
- Facilitates customized campaigns aimed at retaining players.

“With Hadoop on Windows Azure, we can mine data and understand our audience in a way we never could before. It’s really the BI solution for the future.”

Mark Vayman, Lead Program Manager, Halo Services Team

The *Halo* franchise is an award-winning collection of properties that has grown into a global entertainment phenomenon. To date, more than 50 million copies of Halo video games have been sold worldwide. As developers prepared to launch Halo 4, they were tasked with analyzing data to gain insights into player preferences and support an online tournament. To handle those requests, the team used a powerful Microsoft technology called Windows Azure HDInsight Service, based on the Apache Hadoop big data framework. Using HDInsight Service to process and analyze raw data from Windows Azure, the team was able to feed game statistics to the tournament’s operator, which used the data to rank players based on game play. The team also used HDInsight Service to update Halo 4 every week and support a daily email campaign designed to increase player retention. Organizations can also take advantage of data to quickly make business decisions.

Situation

Halo 4 marks the beginning of a new saga in the blockbuster franchise that has shaped entertainment history and defined a generation of gamers. Developed by Microsoft Studios' 343 Industries exclusively for the Microsoft Xbox 360 video game and entertainment system, *Halo 4* brings back the Master Chief character in a new, epic sci-fi adventure. Released in November 2012, the game achieved more than \$220 million in global sales in its first 24 hours and attracted more than 4 million players in its first five days after launch.

For the Halo Services Team, a development team at 343 Industries that manages the game, one of the biggest challenges is scaling to meet player demands. That's one reason the team uses the Windows Azure cloud development platform to power the game's back-end supporting services. These services run the game's key multiplayer features, including leaderboards and avatar rendering. Hosting the multiplayer parts of the game in Windows Azure ensures that the team has a way to quickly and inexpensively add and remove server capacity as needed.

As the game was prepared for release, however, 343 Industries was faced with an entirely new kind of challenge: to gain insight into player behavior and user preferences. To achieve this goal, Microsoft leadership asked 343 Industries to find a way to effectively mine user data.

At the same time, the team was faced with another need: analyzing data during the five-week *Halo 4* "Infinity Challenge" tournament and providing results each day to their tournament partner, Virgin Gaming. The *Halo 4* Infinity Challenge, the largest free-to-enter online *Halo* tournament in the world, tracked a player's personal score in

the game's multiplayer modes across a global leaderboard, giving players a chance to win more than 2,800 prizes. Virgin Gaming needed to use business intelligence (BI) data gathered during the event to update leaderboards on the tournament website.

To meet these business requirements, the 343 Industries knew it needed to find a BI technology solution that would integrate with Windows Azure. "One of the great things about 343 Industries is how they use cutting-edge technology like Azure," says Alex Gregorio, Program Manager for Microsoft Studios, which published *Halo 4*. "So we wanted to find the best BI environment out there, and we needed to make sure it integrated with Azure."

Because all *Halo 4* game data is housed in Azure, the team wanted to find a solution that could effectively produce usable BI information from that data. The team also needed to process this data in the same data center, minimizing storage costs and avoiding charges for data transfers across two data centers. Additionally, the team wanted full control over job priorities, so that the performance and delivery of analytical queries would not be affected by other processing jobs run at the same time. "We had to have a flexible solution that was not on-premises," states Gregorio.

The team began its search for a new BI solution in the months leading up to the scheduled November launch of *Halo 4*.

Solution

Although it initially considered building its own custom BI solution, 343 Industries ultimately decided to use HDInsight Service, which is based on Apache Hadoop, an open-source software framework created by Yahoo! Hadoop, which is ideal for running complex analytics, can analyze

massive amounts of unstructured data in a distributed manner. HDInsight Service is a big data solution for Windows Azure that empowers users to gain new insights from unstructured data, while connecting that data to familiar Microsoft BI tools. "Even though we knew we would be one of the earliest customers of HDInsight Service, it met all our requirements," says Tamir Melamed, Development Manager on the Halo Services Team. "It can run any possible queries, and it is the best format for integration with Azure."

The team was particularly attracted to the flexibility of HDInsight Service, which allowed for separating the amount of the raw data from the processing size needed to consume that data. "With previous systems, we never had the separation between production and raw data, so there was always the question of how running analytics would affect production," says Mark Vayman, Lead Program Manager for the Halo Services Team. "Hadoop solved that problem."

HDInsight Service was also instrumental in changing the team's focus from data storage to useful data analysis. That's because Hadoop applies structure to data when it's consumed, as opposed to traditional data warehouse applications that structure data before it is placed into a BI system.

The team wrote Azure-based services that convert raw game data collected in Azure into the Avro format, which is supported by Hadoop. This data is then pushed from the Azure services in the Avro format into Windows Azure binary large object (BLOB) storage, which HDInsight Service is able to utilize with the ASV protocol. The data can then be accessed by anyone with the right permissions from Windows Azure.

Every day, Hadoop handles millions of data-rich objects related to Halo 4, including preferred game modes, game length, and many other items. With Microsoft SQL Server PowerPivot for SharePoint as a front-end presentation layer, Azure BLOBs are created based on queries from the Halo 4 team.

Microsoft SQL Server PowerPivot for Excel loads data from HDInsight Service using the Hadoop Hive ODBC driver. A PowerPivot workbook is then uploaded to PowerPivot for SharePoint and refreshed nightly within SharePoint, using the connection string stored in the workbook via the Hive ODBC driver to HDInsight Service. The team uses the workbooks to generate reports and facilitate their viewing of interactive data dashboards.

Benefits

Using HDInsight Service, 343 Industries is more agile and can respond faster to customer requests. With the solution's flexibility, the Halo Services Team is able to make weekly updates to the game and was able to help Virgin Gaming detect cheaters in the online Infinity Challenge tournament. HDInsight Service also supports customized email campaigns that the Halo marketing team is using to improve the user experience and retain players. In addition, the solution relies on familiar tools that can be used to simplify decision making.

Increases Agility and Speeds Response Time

With HDInsight Service, 343 Industries is more agile and can respond more quickly to business requests for BI. Part of the reason for this agility is the solution's performance. With Hadoop, the team was able to build a configuration system that can be used to turn various Azure data feeds on or off as needed. "That really helps us get optimal performance, and it's a

big advantage because we can use the same Azure data source to run compute for HDInsight Service on multiple clusters," says Vayman. "It made it easy for us to drive business requests for analysis through an ad-hoc Hadoop cluster without affecting the jobs being run."

And launching Hadoop clusters is a simple, fast process. "We can easily launch a new Hadoop cluster in minutes, run a query, and get back to the business in a few hours or less," says Melamed. "Azure is very agile by nature, and Hadoop on Azure is more powerful as a result."

Helps Halo 4 Team Make Weekly Game Updates

In addition to responding quickly to business requests, the Halo 4 team can take BI data pulled from the game each day and identify user trends, such as the average length of a game and the specific game features that players use the most. By getting these insights, the Halo 4 team can make frequent updates to the game. "Based on the user preference data we're getting from Hadoop, we're able to update game maps and game modes on a week-to-week basis," says Vayman. "And the suggestions we get in the forums often find their way into the next week's update. We can actually use this feedback to make changes and see if we attract new players. Hadoop and the forums are great tuning mechanisms for us."

The team is also taking user feedback and giving it to the game's designers, who can consider the suggestions in developing future editions of Halo.

Provides In-Game Analysis and Helps Identify Cheaters

Because Hadoop applies structure to data when it's consumed, the team can focus more on analytics and less on storage.

Instead of worrying about how to store and structure game data, the team can concentrate on what game modes users play in or how many users are playing at any given time. With this ability to focus more tightly on analysis, 343 Industries could meet the needs of Virgin Gaming. "Using Microsoft HDInsight Service, we were able to analyze the data during the five weeks of the Halo 4 Infinity Challenge," says Vayman. "With the fast performance we got from the solution, we could feed that data to Virgin Gaming so it could update the leaderboards on the tournament website every day."

In addition, because of the way the team set up Hadoop to work within Azure, the team was able to detect cheaters during the Halo 4 Infinity Challenge. "HDInsight Service gives us the ability to easily read the data," says Vayman. "In this case, there are many ways in which players try to gain extra points in games, and we could look back at previous data stored in Azure and identify user patterns that fit certain cheating characteristics, which was unexpected." After receiving this data from the team, Virgin Gaming sent out a notification that any player found or suspected of cheating would be removed from the leaderboards and the tournament.

Contributes to Player Retention

The flexibility of the HDInsight Service BI solution also gives 343 Industries a way to reach out to players through customized campaigns, such as the series of email blasts the team sent to players immediately after the launch. For that campaign, the team set up Hadoop queries to identify users who started playing on a certain date. The team then wrote a file and placed it into a storage account on Windows Azure, where it was sent through SQL Server 2008 R2 Integration Services into a database owned by the Xbox marketing team.

The marketing team then used this data to send these new players emails customized by screening several variables including when they started playing Halo 4 and their game play behaviors. The choice of which email each player received was determined by the HDInsight Service system. "That gave marketing a new way to retain users and keep them interested by talking about new aspects of the game," Gregorio says. The Halo marketing team plans to run similar email campaigns for the game until a new edition is released. "Basing an email campaign on HDInsight Service and Hadoop was a big win for the marketing team, and also for us," adds Vayman. "It showed us that we were able to use data from HDInsight Service to customize emails, and to actually use BI to improve the player experience and affect game sales."

Uses Familiar Tools to Simplify Decision Making

Microsoft has started to expand HDInsight Service to other internal groups, and one of the reasons adoption is growing is that users do not have to be engineers or Hadoop experts to take advantage of the technology. Data is collected in Azure and made easily accessible through familiar productivity tools. "By hooking Hadoop into a set of tools that are already familiar, such as Microsoft Excel or Microsoft SharePoint, people can take advantage of the power of Hadoop without needing to know the technical ins and outs," says Vayman. "A good example of that is the data about Halo 4 Infinity Challenge cheaters that we gave to Virgin Gaming. The people receiving that data are not Hadoop experts, but they can still easily use the data to make business decisions."

Another reason Hadoop is becoming more widely used is that the technology

continues to evolve into an increasingly powerful tool. "The traditional role of BI within Hadoop is expanding because of the raw capabilities of the platform," says Brad Sarsfield, Microsoft SQL Server Developer. "In addition to just BI reporting, we've been able to add predictive analytics, semantic indexing, and pattern classification, which can all be leveraged by the teams using Hadoop."

With these and other capabilities, there is little question that HDInsight Service will continue to positively affect business. "With Hadoop on Windows Azure, we can mine data and understand our audience in a way we never could before," says Vayman. "It's really the BI solution for the future."

For More Information

For more information about Microsoft products and services, call the Microsoft Sales Information Center at (800) 426-9400. In Canada, call the Microsoft Canada Information Centre at (877) 568-2495. Customers in the United States and Canada who are deaf or hard-of-hearing can reach Microsoft text telephone (TTY/TDD) services at (800) 892-5234. Outside the 50 United States and Canada, please contact your local Microsoft subsidiary. To access information using the World Wide Web, go to:
www.microsoft.com

Windows Azure

With Windows Azure, you can quickly build, deploy, and manage applications across a global network of Microsoft-managed data centers. You can build applications using any operating system, language, or tool.

[For more information, please visit
www.windowsazure.com.](http://www.windowsazure.com)

Software and Services

- Microsoft Server Product Portfolio
 - Microsoft SQL Server 2008 R2 Enterprise
 - Windows Azure
 - Microsoft SQL Server PowerPivot for SharePoint
 - Windows Azure HDInsight Services

Technologies

- Apache Hadoop
- Microsoft SQL Server PowerPivot for Excel