Unifi Software benefits from unified approach to data preparation and management

MATT ASLETT

30 DEC 2016

The company has more than doubled its roster of paying customers since the start of the year, benefitting from its combined focus on data management as an enabler of self-service data preparation.

451 Research®

When we initiated coverage of Unifi Software earlier this year, we noted that it was well positioned to take advantage of enterprises' emerging plans to create multifunction 'data lake' architectures with its approach of addressing the needs of both business users and IT administrators in a single product. While it is still early stages, the company has more than doubled its number of paying customers since then and has expanded both its sales and development capabilities.

THE 451 TAKE

We previously noted that we expected Unifi Software to flourish given its focus on data lake management and self-service data preparation. While some of the self-service data-preparation specialists can boast larger installed bases thanks to their focus on data analysts and business users, Unifi and other data lake management providers are adopting a more strategic approach to also address the needs of data stewards and IT professionals. We believe addressing the requirements of both sets of users will be essential in delivering 'managed self-service' environments that offer the benefits of both self-service as well as data access and governance controls. Unifi notes that it continues to see some resistance from IT professionals in moving toward greater data democratization. However, it is also seeing more and more enterprises embracing the need to be data-driven – thus, its approach of addressing the requirements of those IT professionals is likely to stand it in good stead in the long term.

CONTEXT

The importance of a combination of self-service data preparation and management as a driver for data lake architectures was a major focus for 451 Research's Data Platforms and Analytics channel in 2016. Back in March, we argued that the combination of governance and preparation was fundamental to the creation of functioning data lakes, while in our recent look at the trends likely to shape the market in 2017 we examined the movement toward managed self-service environments that balance the benefits of empowering data analysts and scientists to prepare and analyze their own data sets with the requirements of data stewards and IT to maintain control over who has access to what data, and for what purpose.

This trend is especially significant for Unifi Software, which was founded in 2014 by former executives from analytics database firm Greenplum and advanced analytics software provider Alpine Data to develop its catalog-based approach to data integration and preparation. While there are numerous vendors offering self-service data-preparation software, as well as catalog-based approaches to data lake management, Unifi is differentiated (although not unique – see Competition, below) in attempting to address the needs of both business analysts and IT professionals with a single offering.

The company describes its approach as providing self-service data preparation that is deployed by business users and managed by IT professionals, and it's a strategy that appears to be paying off. When we were introduced to Unifi in March, it had 10 paying customers, including Havas Media and A.T. Kearny. It recently hit its target of reaching 25 paying clients before year-end, and its list of reference customers now includes Sabre Holdings, Boston Biomedical, Compassion-First Pet Hospitals and Essentra. It remains early stages for the deployment of self-service data preparation and catalog-based data lake management, however, and the company concedes that it is not unusual to face resistance from IT departments that have more traditional approaches to data management and governance that essentially focus on restricting access to data.

Unifi's argument is that in order to transition to becoming more data-driven, enterprises must focus on data management and governance as enablers of self-service data preparation, democratizing access to data while maintaining access controls for privacy, security, regulatory and data-quality reasons. This is a strategy that an increasing number of organizations are beginning to embrace, however. The company points to an example of a multinational alcoholic beverage vendor (that cannot currently be named) that has combined sales, marketing and supply chain data to better track kegs of beer in its distribution chain and initiate promotions to maximize sales and avoid wastage of product unsold within its eight-week quality lifespan. The solution to its problem was only identified by opening up access to those multiple data sources internally, notes Unifi.

The company's software is designed to enable business users to integrate data from multiple data sources – including structured and unstructured data from relational and non-relational databases and Hadoop, as well as files, data services and social media – into a single cataloged data set on Hadoop. Key to this is the company's OneParse capability, which automatically identifies whether data is structured or semi-/unstructured and normalizes the data to represent it to business users in a tabular format.

OneParse creates a metadata catalog of the combined data set, while Unifi also provides end users with a simple natural language search-based interface to discover data, and IT professionals with LDAP-compatible access controls and data-authorization capabilities to implement data privacy and policy requirements – for example, to limit access to PII data. Additionally, the company offers IT professionals data lineage, resource management, workflow and scheduling capabilities, and business users an interface for integrating, transforming, cleansing and visualizing data, as well as collaborative capabilities. Unifi also takes advantage of underlying machine-learning functionality to provide recommendations to business users on potential transformation tasks based on the nature of the data.

The company has expanded its headcount since we last spoke, officially adding a team of developers in Bangalore that helped boost its headcount to 43, from 30 or so in March. Another recent hire is VP of sales Chris Flynn – previously at Zoomdata, as well as Greenplum, ParAccel and Opsware – who is leading Unifi's focus on Fortune 1000 firms. The company continues to be led by its cofounders: CEO Rob Carlson, VP of business development Jim Carr, VP of product Sean Keenan and VP of engineering Ayush Parashar.

Unifi has raised $14.5m in funding, including a $4.5m seed round in 2014 led by Canaan Partners and also including Omaha Capital. The company subsequently landed a $10m series A round in late 2015 led by Pelion Venture Partners, with participation from Canaan Partners.

COMPETITION

As noted, there are numerous vendors targeting both self-service data preparation and data lake management, although few are doing so with a combined offering. We believe Unifi Software's most directly comparable competitor is Podium Data, which is similarly focused on 'strategic' adoption driven by data stewards and IT professionals, but while also addressing the self-service data-preparation needs of business users and analysts. Zaloni also targets both business users and IT professionals but offers two separate products: Mica for self-service data management, and Bedrock Data Management for data lake management. Meanwhile, Alation has refined its focus on data stewards and IT professionals with its Data Catalog as an enabler for self-service analytics, although many deals continue to be driven by analytics users and senior analytics leaders.

Unifi also encounters self-service data-preparation specialists such as Trifacta and Paxata, and although the former targets business users almost exclusively, the latter is paying closer attention to the requirements of IT professionals with its plans to deliver semantic data catalog capabilities. Today, however, Unifi is more likely to encounter Trifacta and Paxata in combination with a data lake management provider such as Waterline Data or Tamr. Other players that have entered this space in recent months and years include data analytics firms like Alteryx, Datawatch, TIBCO and SAS Institute, as well as data management suppliers such as IBM, Talend, Informatica, Hitachi's Pentaho and Oracle. Syncsort also recently entered this space via its acquisition of Trillium Software from Harte Hanks, although the target's Prepare and Refine products are based on Unifi.

SWOT ANALYSIS

STRENGTHS
Unifi Software was founded by a management team with data management and analytics experience and has raised a healthy amount of early-stage funding.

WEAKNESSES
Some of the early self-service data-preparation vendors have the advantage of more viral user-driver adoption approaches, although Unifi can point to a broader set of functionality to address IT requirements.

OPPORTUNITIES
We are noting a growing focus on data management and governance as an enabler of self-service data preparation and data lakes, which is playing to Unifi's strengths.

THREATS
The incumbent data management and integration vendors already have supplier relationships with established enterprises and will look to squeeze out emerging data lake specialists.