Text 1

Hobit

V jisté podzemní noře bydlel jeden hobit. Nebyla to žádná ošklivá, špinavá, vlhká díra, plná konečků žížal a páchnoucí slizem, ani vyschlá, holá písčitá jáma, kde by se nedalo na nic sednout a ničeho se najíst: byla to hobití nora, a to znamená komfort. Měla dokonale okrouhlé dveře jako okénko lodní kajuty, nalakované nazeleno a s nablýskaným žlutým knoflíkem mosazné kliky předně uprostřed. Dveře vedly do předsíně jako do válcovitého tunelu, velmi komfortního tunelu bez kouře, se stěnami obloženými dřevěným táflováním, a s kachlíčkovanou podlahou pokrytou koberci, zařízeného leštěnými křesly a velikou spousto věšáků na kloubouky a kabáty, poněvadž hobit rád přijímal návštěvy. Tune se vinul hezký kousek cesty dost hluboko, ale ne docela přím do úbočí kopce – Kopce s velkým K, jak mu říkali všichni obyvatelé na míle kolem, -a z tunelu vedlo mnoho okrouhlých dvířek napřed na jednu stran a potom na duhu. Hobit si nepotrpěl na žádné chození do schodů: ložnice, koupelny, sklepy, spižírny (a těch měl spoustu), šatny (měl celé místnosti vyhrazené oblečení), kuchyně, jídelny, to všechno bylo v témž patře, a vlastně na téže chodbě. Nejlepší pokoje ležely všechny po levé staně (ve směru příchoduú, protože jedině ty měl okna, hluboce zasazená okrouhlá okna,…

Autor: J.R.R.Tolkien

Název knihy: Hobit

Nakladatelství: Odeon

Rok vydání: 1978

Text 2

This article is an excerpt from our comprehensive, 40-page eBook:

The Architect’s Guide to Streaming Data and Data Lakes. Read on to discover design patterns and guidelines for for streaming data architecture, or get the full eBook now (FREE) for in-depth tool comparisons, case studies, and a ton of additional information.

Streaming data is becoming a core component of enterprise data architecture due to the explosive growth of data from non-traditional sources such as IoT sensors, security logs and web applications.

Streaming technologies are not new, but they have considerably matured in recent years. The industry is moving from painstaking integration of open-source Spark/Hadoop frameworks, towards full stack solutions that provide an end-to-end streaming data architecture built on the scalability of cloud data lakes.

able of Contents

What is Streaming Data and Streaming data Architecture?

Why Streaming Data Architecture? Benefits of Stream Processing

The Components of a Streaming Architecture

Modern Streaming Architecture

The Future of Streaming Data in 2019 and Beyond

Further reading

What is Streaming Data and Streaming data Architecture?

Streaming data refers to data that is continuously generated, usually in high volumes and at high velocity. A streaming data source would typically consist of a stream of logs that record events as they happen – such as a user clicking on a link in a web page, or a sensor reporting the current temperature.

Common examples of streaming data include:

IoT sensors

Server and security logs

Real-time advertising

Click-stream data from apps and websites

In all of these cases we have end devices that are continuously generating thousands or millions of records, forming a data stream – unstructured or semi-structured form, most commonly JSON or XML key-value pairs. Here’s an example of how a single streaming event would look – in this case the data we are looking at is a website session (extracted using Upsolver’s Google Analytics connector):

A single streaming source will generate massive amounts of these events every minute. In its raw form, this data is very difficult to work with as the lack of schema and structure makes it difficult to query with SQL-based analytic tools; instead, data needs to be processed, parsed and structured before any serious analysis can be done.

Streaming Data Architecture

A streaming data architecture is a framework of software components built to ingest and process large volumes of streaming data from multiple sources. While traditional data solutions focused on writing and reading data in batches, a streaming data architecture consumes data immediately as it is generated, persists it to storage, and may include various additional components per use case – such as tools for real-time processing, data manipulation and analytics.

Streaming architectures need to be able to account for the unique characteristics of data streams, which tend to generate massive amounts of data (terabytes to petabytes) that it is at best semi-structured and requires significant pre-processing and ETL to become useful.

Stream processing is a complex challenge rarely solved with a single database or ETL tool – hence the need to ‘architect’ a solution consisting of multiple building blocks. Part of the thinking behind Upsolver is that many of these building blocks can be combined and replaced with declarative functions within the platform, and we will demonstrate how this approach manifests within each part of the streaming data supply chain.

Why Streaming Data Architecture? Benefits of Stream Processing

Stream processing used to be a ‘niche’ technology used only by a small subset of companies. However, with the rapid growth of SaaS, IoT and machine learning, organizations across industries are now dipping their feet into streaming analytics. It’s difficult to find a modern company that doesn’t have an app or a website; as traffic to these digital assets grows, and with increasing appetite for complex and real-time analytics, the need to adopt modern data infrastructure is quickly becoming mainstream.

While traditional batch architectures can be sufficient at smaller scales, stream processing provides several benefits that other data platforms cannot:

Able to deal with never-ending streams of events—some data is naturally structured this way. Traditional batch processing tools require stopping the stream of events, capturing batches of data and combining the batches to draw overall conclusions. In stream processing, while it is challenging to combine and capture data from multiple streams, it lets you derive immediate insights from large volumes of streaming data.

Real-time or near-real-time processing—most organizations adopt stream processing to enable real time data analytics. While real time analytics is also possible with high performance database systems, often the data lends itself to a stream processing model.

Detecting patterns in time-series data—detecting patterns over time, for example looking for trends in website traffic data, requires data to be continuously processed and analyzed. Batch processing makes this more difficult because it breaks data into batches, meaning some events are broken across two or more batches.

Easy data scalability—growing data volumes can break a batch processing system, requiring you to provision more resources or modify the architecture. Modern stream processing infrastructure is hyper-scalable, able to deal with Gigabytes of data per second with a single stream processor. This allows you to easily deal with growing data volumes without infrastructure changes.

To learn more, you can read our previous article on stream vs batch processing.



Text 3:

Osnova Přednášky: 1) Informatika – Informační technologie, data, informace, systém, řízení a procesy. 2) Portfolio, infrastruktura a funkcionalita aplikací. 3) Kancelářský software – pokročilá úroveň MS Excel a MS Word. 4) Úvod do programovacích a databázových jazyků. 5) Programové a databázové jazyky. 6) Transakční aplikace – funkcionalita ERP produktů. 7) Aplikace pro podporu rozhodování a pro řízení externích vztahů. 8) Aplikace zaměřené na podporu produktivity práce. 9) Řízení provozu informatiky v podniku. 10) Řízení informatických podnikových procesů. 11) Životní cyklus aplikací. 12) Bezpečnost v informatice. 13) Průmysl 4.0. Semináře: 1) Práce s textem - základní funkce MS Word. 2) Práce s textem - pokročilé funkce MS Word. 3) Prezentační techniky - pokročilé funkce MS Powerpoint. 4) Práce s tabulkovými procesory I. - MS Excel, základní úroveň, rovnice, práce s grafy, formátování, textové, logické funkce a datum a čas. 5) Práce s tabulkovými procesory II. - MS Excel, pokročilá úroveň, kontingenční tabulky, matematické a Ing. Květa Papoušková, učo 23056 statistické. 6) Práce s tabulkovými procesory III. - MS Excel, pokročilá úroveň, vyhledávácí, databáze. 7) Základy programování - program a programovací jazyky, základní algoritmy a psaní kódu v PHP. 8) Práce s databázemi - data a databáze, základní dotazy typu SQL na příkladu MySQLi, základní dotazy v prostředí programového kódu v PHP. 9) Programování – jednoduchá aplikace v PHP. 10) Aplikace pro ekonomy I - používání cloudových služeb Google Apps, Office 365. 11) Aplikace pro ekonomy II - používání ERP systémů. 12) Aplikace pro ekonomy III - používání nástrojů evidence. 13) Aplikace pro ekonomy IV - hodnocení a plánování zaměstnanců. Literatura povinná literatura SKLAR, D., 2018. PHP 7: praktický průvodce nejrozšířenějším skriptovacím jazykem pro web. Brno: Zoner Press. ISBN 978-80-7413-363- 3. TOPOLOVÁ, I., T. KUBÁLEK a M. KUBÁLKOVÁ, 2017. Textový procesor Microsoft Word 2016. [Praha]: Oeconomica, nakladatelství VŠE. ISBN 978-80-245-2198-5. BORONCZYK, T., 2016. MySQL okamžitě. Brno: Computer Press. ISBN 978-80-251-4737-5. KUBÁLKOVÁ, M., T. KUBÁLEK a I. TOPOLOVÁ, 2017. Tabulkový program Microsoft Excel 2016. Praha: Oeconomica, nakladatelství VŠE. ISBN 978-80-245-2229-6. DASGUPTA, S., 2016. Computer Science: A Very Short Introduction. New York: OUP Oxford. ISBN 978- 0198733461. GÁLA, L., J. POUR a Z. ŠEDIVÁ, 2015. Podniková informatika: počítačové aplikace v podnikové a mezipodnikové praxi. 3., aktualiz. vyd. Praha: Grada. ISBN 978-80-247-5457-4. doporučená literatura WELLING, L. a L. THOMSON, 2017. Mistrovství PHP a MySQL. Brno: Computer Press. ISBN 978-80- 251-4892-1. PAVLÍČEK, A., A. GALBA a M. HORA, 2017. Moderní informatika. Druhé, rozš. vyd. [Praha]: Professional Publishing. ISBN 978-80-906594- 6-9. NAVRÁTIL, P., 2015. Příklady a cvičení z informatiky: zadání. 3. vyd. Prostějov: Computer Media. ISBN 978-80-7402-160-2. FEREBAUEROVÁ, R. a O. PEKÁREK, 2014. Aplikovaná informatika. České Budějovice: Vysoká škola evropských a regionálních studií. ISBN 978-80-87472-74-3. WELLING, L. and L. THOMSON, 2016. PHP and MySQL Web Development. 5th edit. Addison Wesley, Hoboken, NJ. TOPOLOVÁ, I., T. KUBÁLEK a M. KUBÁLKOVÁ, 2016. Prezentační program Microsoft PowerPoint 2016. Praha: Oeconomica, nakladatelství VŠE. ISBN 978-80-245-2138-1.

Zadání: Upravte první dva texty dle zadání v druhém souboru. Třetí text sslouží k domácí přípravě a tríninku.