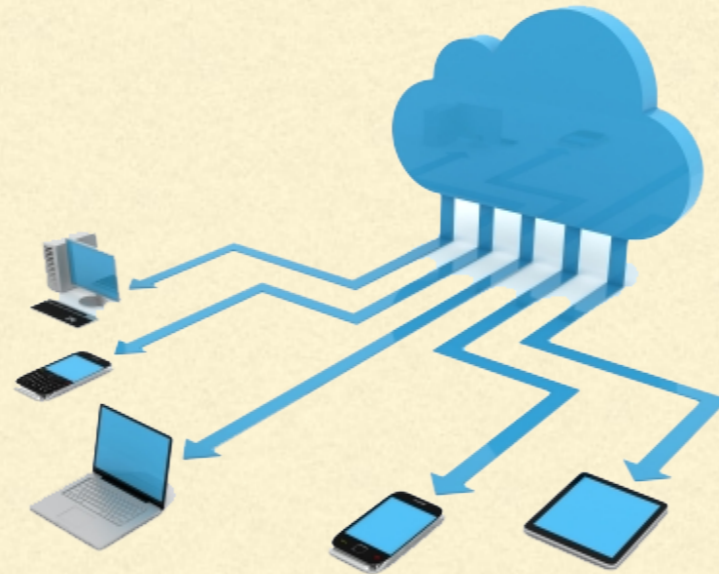




VIRTUALIZACE A CLOUD

Jaroslav Žáček

jaroslav.zacek@osu.cz



Cloud

- Self-service
- PaaS & SaaS
- Automated management out of the box
- Empowerment/on-demand computing
- Distributed control
- Made for end users
- Scalability

Virtualization

- Automated management possible with add-ons
- Virtual machines not inherently more secure than in-house hardware
- Licensing model tied to hardware; IT must provision resources
- Manual deployment & retraction
- Centralized control required
- Made for IT

- Infrastructure as a Service
 - Abstraction of hardware from user interface
-

VIRTUALIZACE

- Postupy, techniky a prostředky k přístupu k dostupným zdrojům jiným způsobem, než jak fyzicky existují
 - Vcelku starý termín - 60. léta 20. století
 - **Hypervisor** - Software nebo firmware, který vytváří stroj na hostovaném hardware
 - Emulace x Simulace
 - Emulace - pouze část problému, většinou implementována pomocí stavových automatů
 - Simulace - simuluje celý hardware, např. při tvorbě procesorů
-

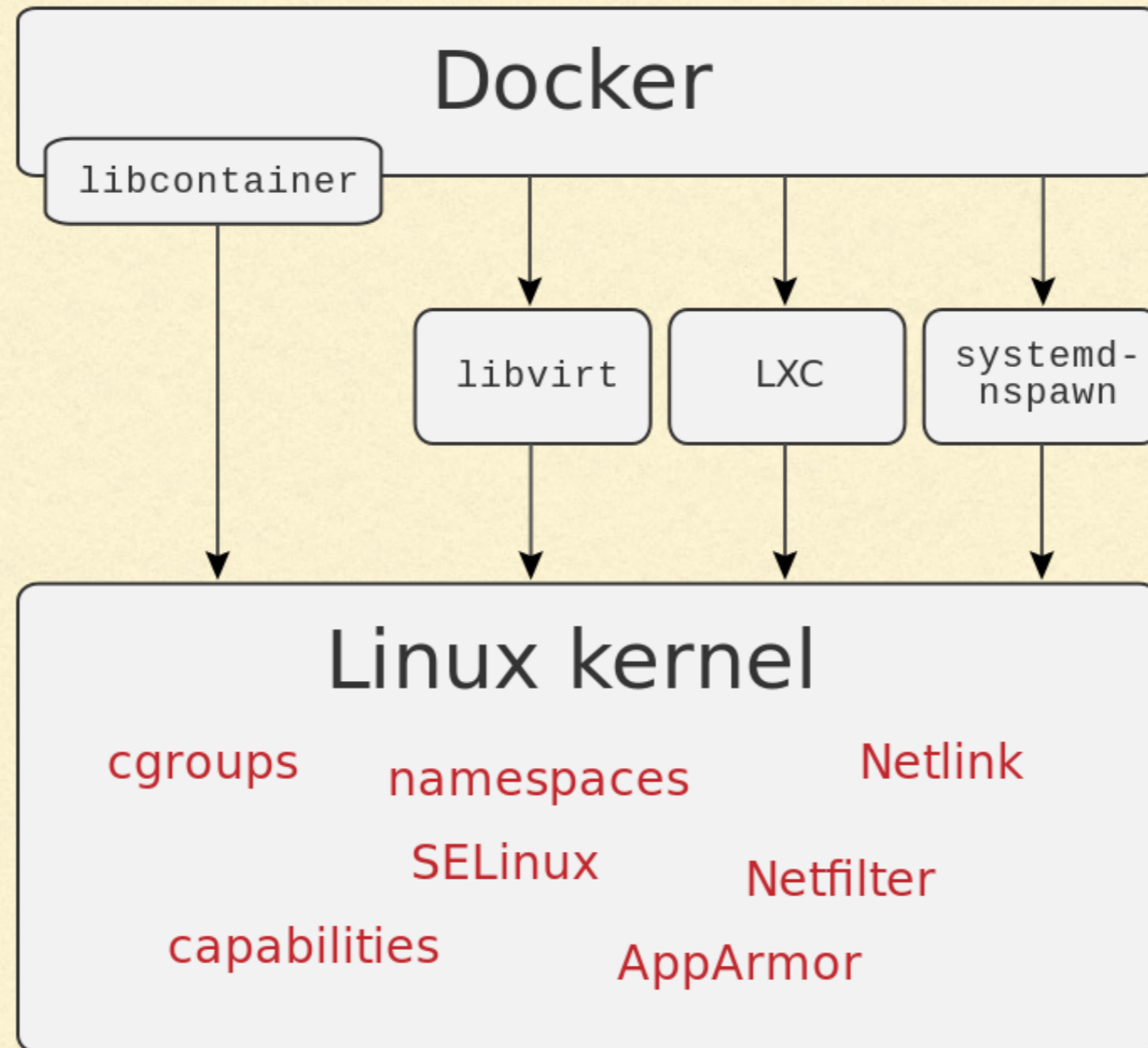
VIRTUALIZACE

- **Nativní** - Simuluje dostatečné množství hardware pro běh neupraveného OS pro stejný druh procesoru.
 - Typičtí zástupci - VMWare Workstation i Server, Parallels Desktop, VirtualBox
 - **Částečná virtualizace** - Virtuální stroj simuluje více instancí mnoha prostředí hardware - podporuje sdílení zdrojů a izolaci procesů.
-

VIRTUALIZACE

- **Paravirtualizace** - Nesimuluje více instancí, ale nabízí konkrétní API, které může být použito z upraveného hostovaného OS. Nutná podpora hardware - volá se speciální instrukcí DIAG.
 - **Virtualizace na úrovni OS** - virtualizuje se fyzický server na úrovni OS.
 - **Aplikační virtualizace** - aplikací používají místní zdroje, ale běží ve zvláštním virtuálním stroji.
-

KAM PATŘÍ DOCKER?



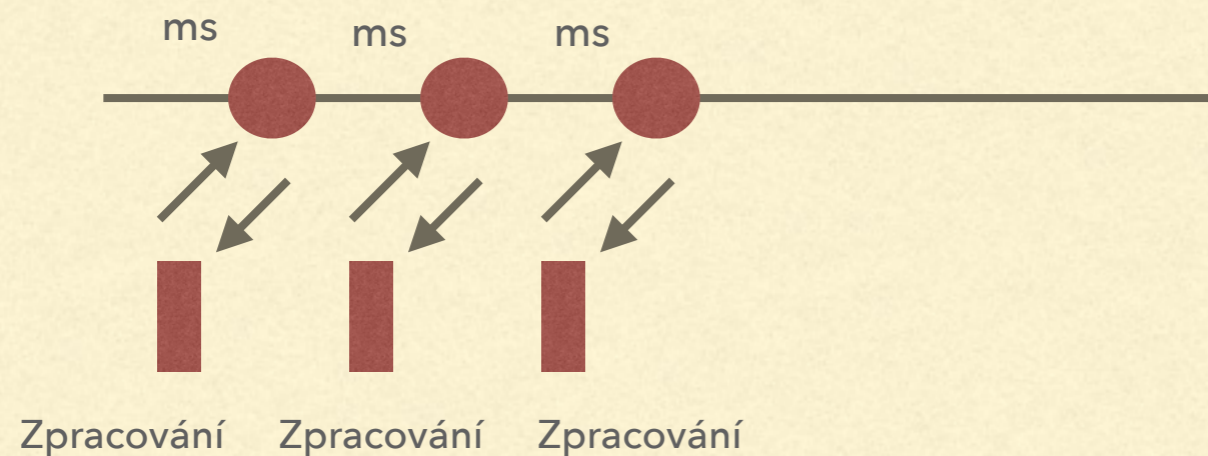
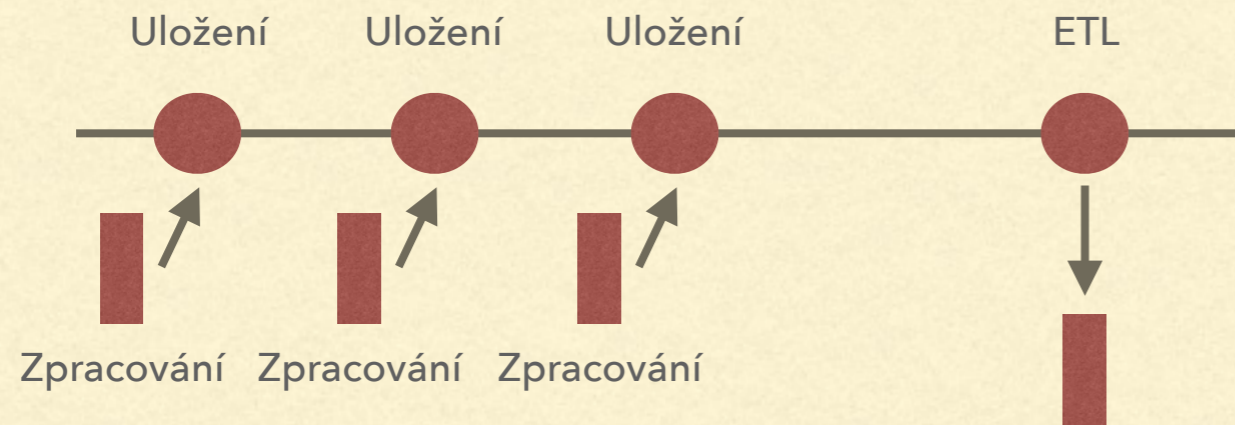
BIG DATA

- Za jednu minutu:
 - Uživatelé YouTube nahrají 300 hodin videa
 - Uživatelé Facebooku "dají Like" 4 166 667 příspěvkům
 - Uživatelé Twitteru vystaví 347 222 příspěvků
-

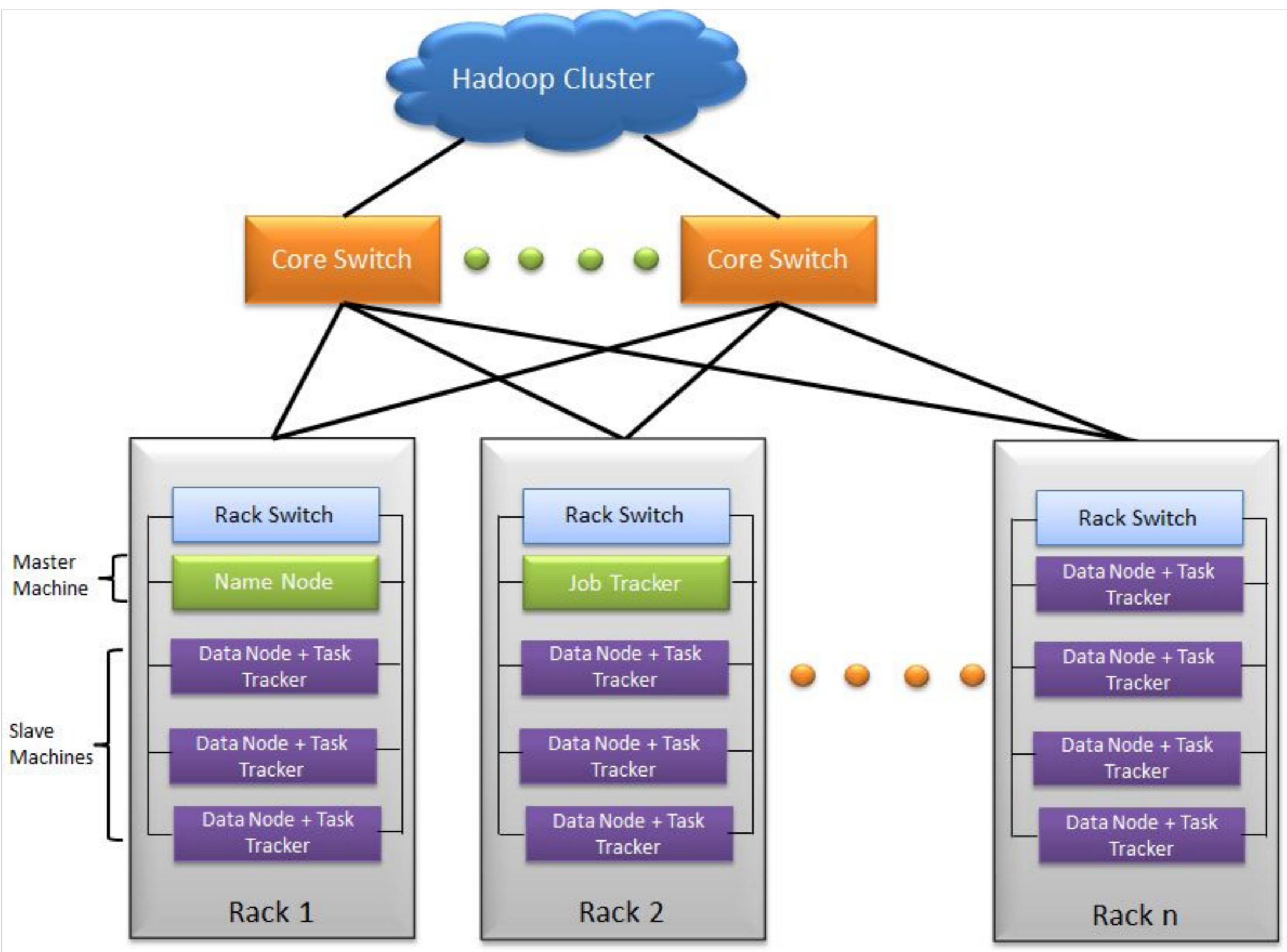
ANALÝZA “VELKÝCH DAT”

- Proces, při kterém procházíme velké kolekce dat pro nalezení skrytých vzorů, korelací, trendů a uživatelských preferencí.
 - Dva typy zpracování:
 - Dávkové
 - V reálném čase
-

DRUHY ZPRACOVÁNÍ

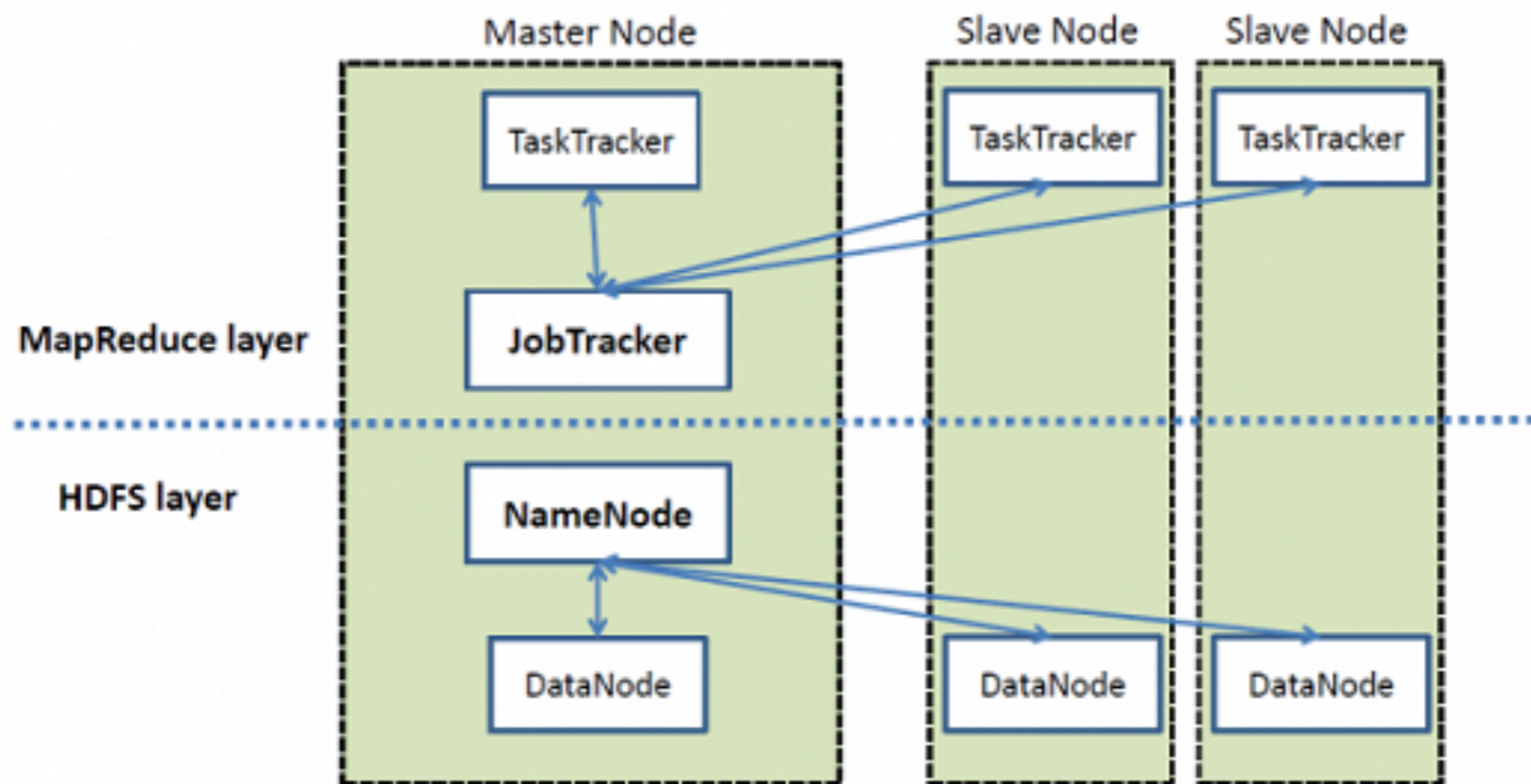


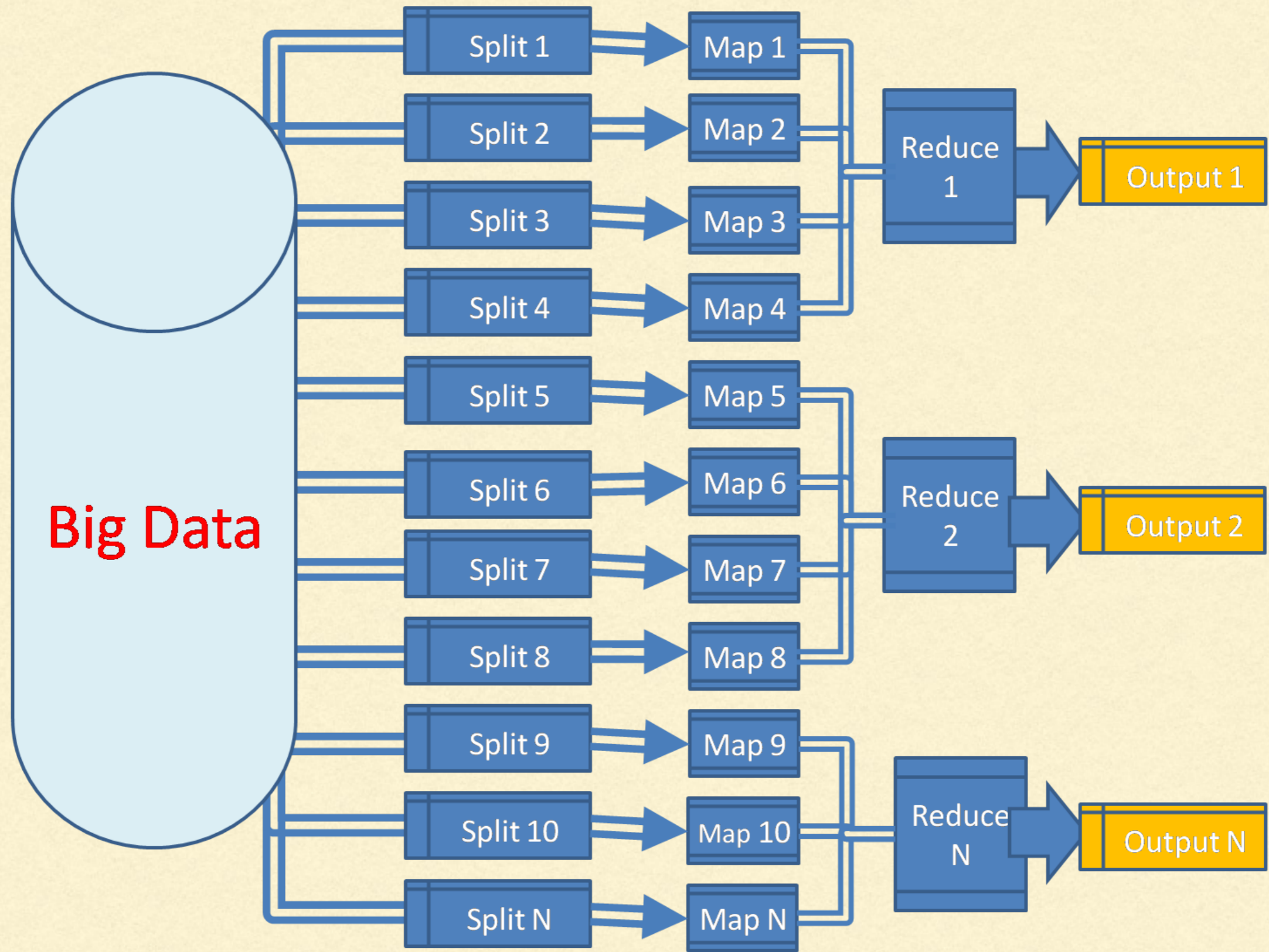
- Data se sbírají určité časové období a zpracovávají v dávce
- Analýza se provádí v reálném čase



APACHE HADOOP

High Level Architecture of Hadoop





MAPREDUCE

Vstup

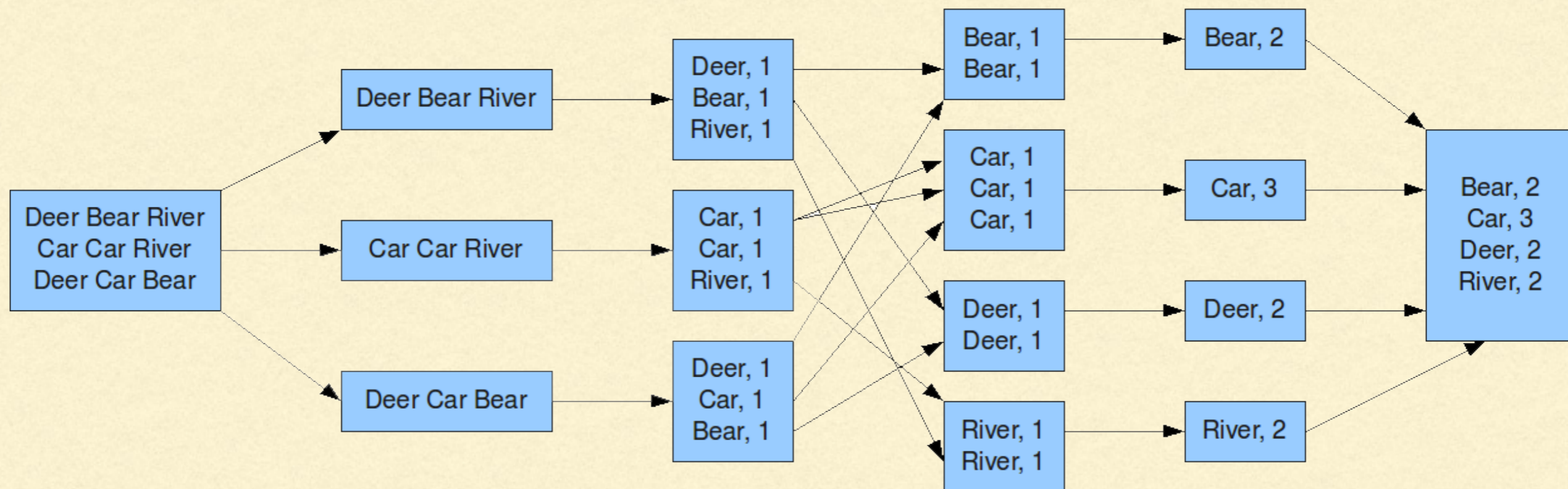
Splitting

Mapping

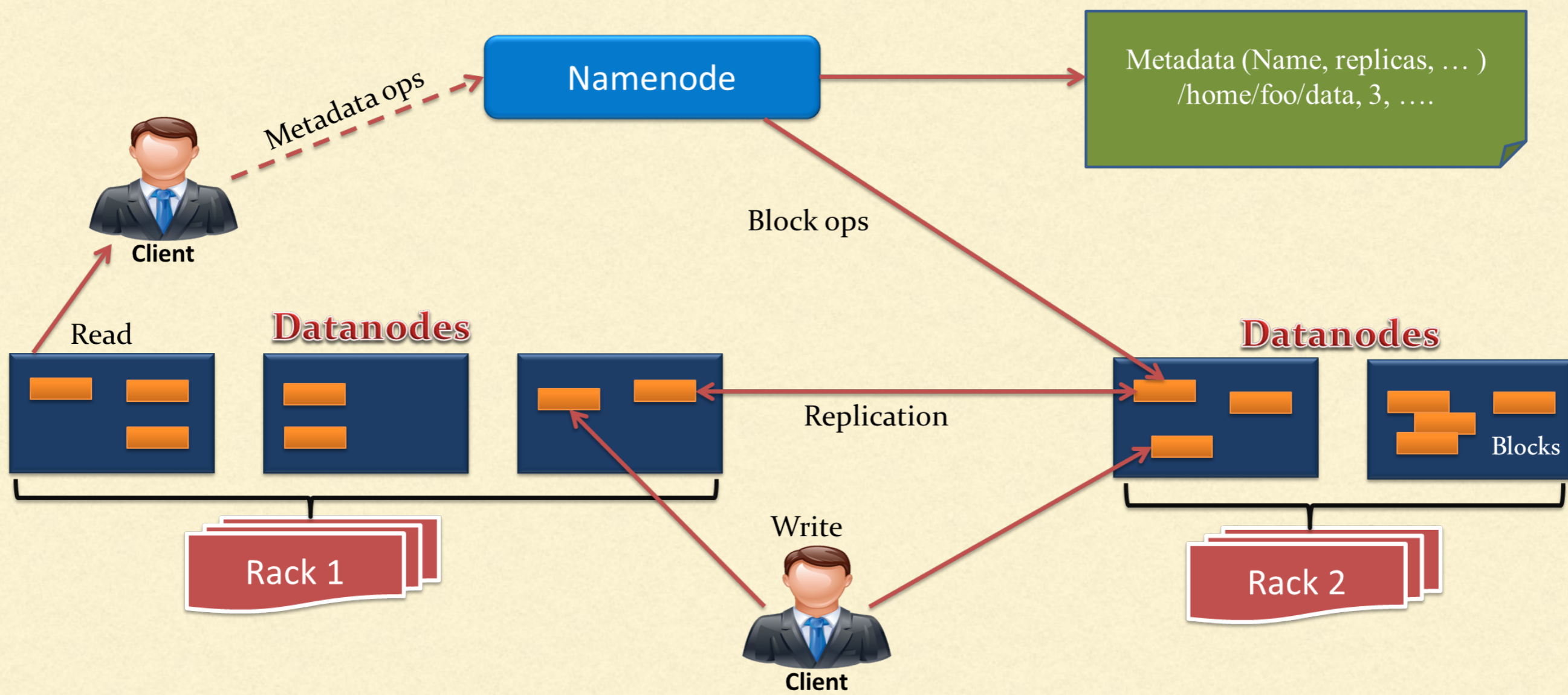
Shuffling

Reducing

Výstup

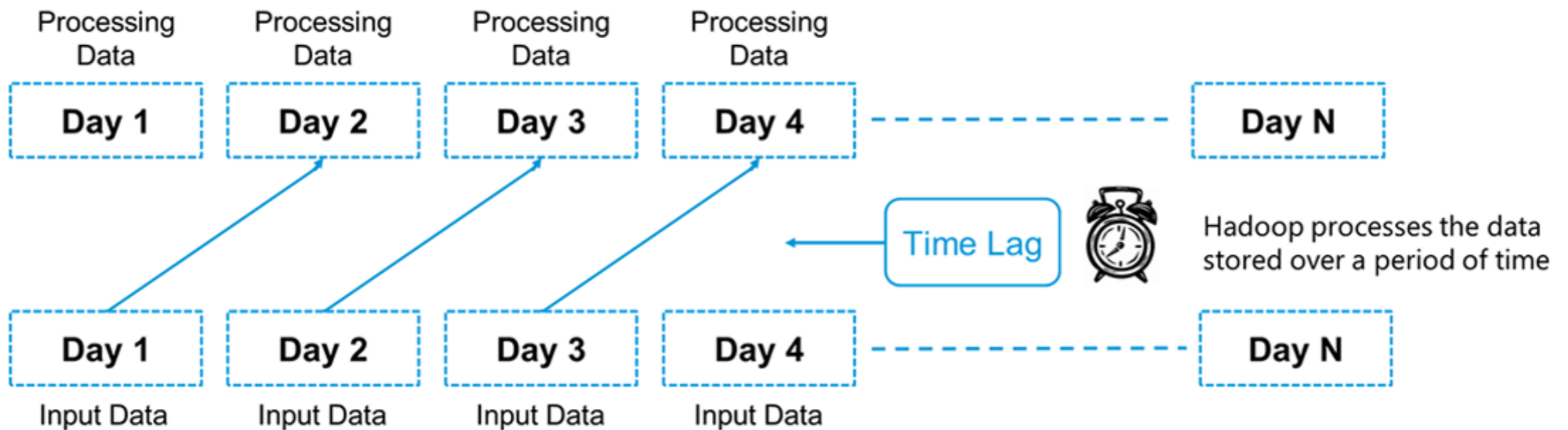


HDFS ARCHITEKTURA





Processing Data Using MapReduce





Real Time Processing in Spark

