Large datasets and data streams in CSC computing environment

Aleksi Kallio

IoT workshop 15.1.2019

Storing large data

Data storage

- Working storage for active data
  - HPC computing environment offers user and project directories for active data as well as an archive
  - Cloud environments provide storage for computing, processing and analyzing data
  - Pouta Object Storage is a cross-platform service for storing and sharing data
  - Databases for research are available for computing projects

- Services for storing stable data
  - IDA for safe storing and sharing

- More information: http://research.csc.fi/storage

Processing large datasets and streams

Computing for large data

- HPC
  - Taito cluster
  - Sisu supercomputer
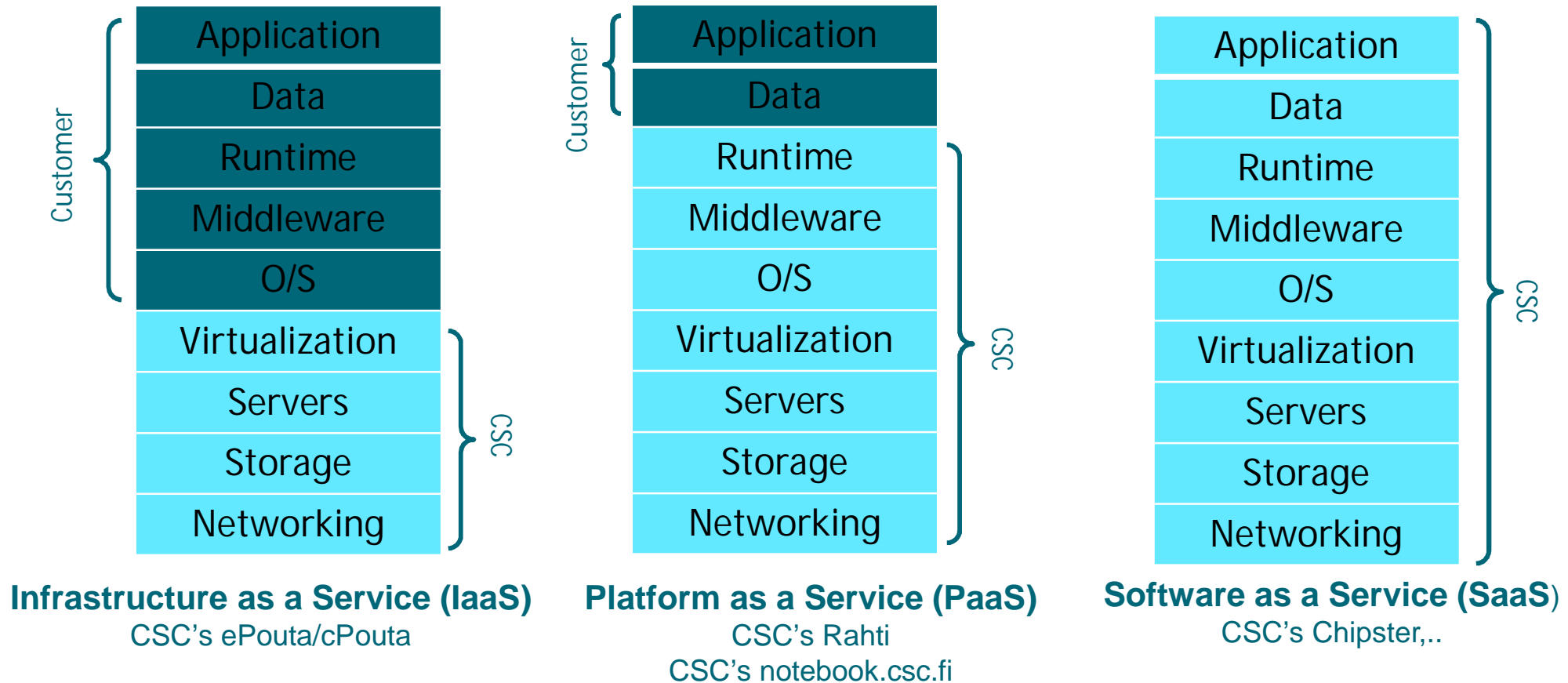
=> Scientific custom streaming applications

- Cloud
  - Pouta (virtual machines / OpenStack)
  - Rahti (containers / OpenShift / Kubernetes)

=> Standard tools Kafka, Spark Streaming, Storm…

- More information: https://research.csc.fi/computing-and-software

Cloud Service Models offered by CSC

Infrastructure as a Service (IaaS)
CSC's ePouta/cPouta

Platform as a Service (PaaS)
CSC's Rahti
CSC's notebook.csc.fi

Software as a Service (SaaS)
CSC's Chipster,..

Typical Resources You get from Pouta Clouds

CSC

- VMs
- Oversubscribed or dedicated CPUs
- GPUs

Compute

- Volume Storage
- Object Storage

Storage

- 10 GbE or 40 GbE

Private VLAN

- With or Without NAT

IPv4

- With Latest Security patches

Images

OS

- Full programmability of your resources

API

Pouta: Hardware Options

- AI, Deep Machine Learning
- Intensive 3D/Video/Image processing

GPU Flavor

- Hadoop/Spark
- Clustered Databases

I/O Flavor

Standard Flavor

- Normal Computing & Software Development
- Normal Application/ Web & content services hosting

HPC Flavor

- Scientific Applications
- Advanced Computing & Software Development
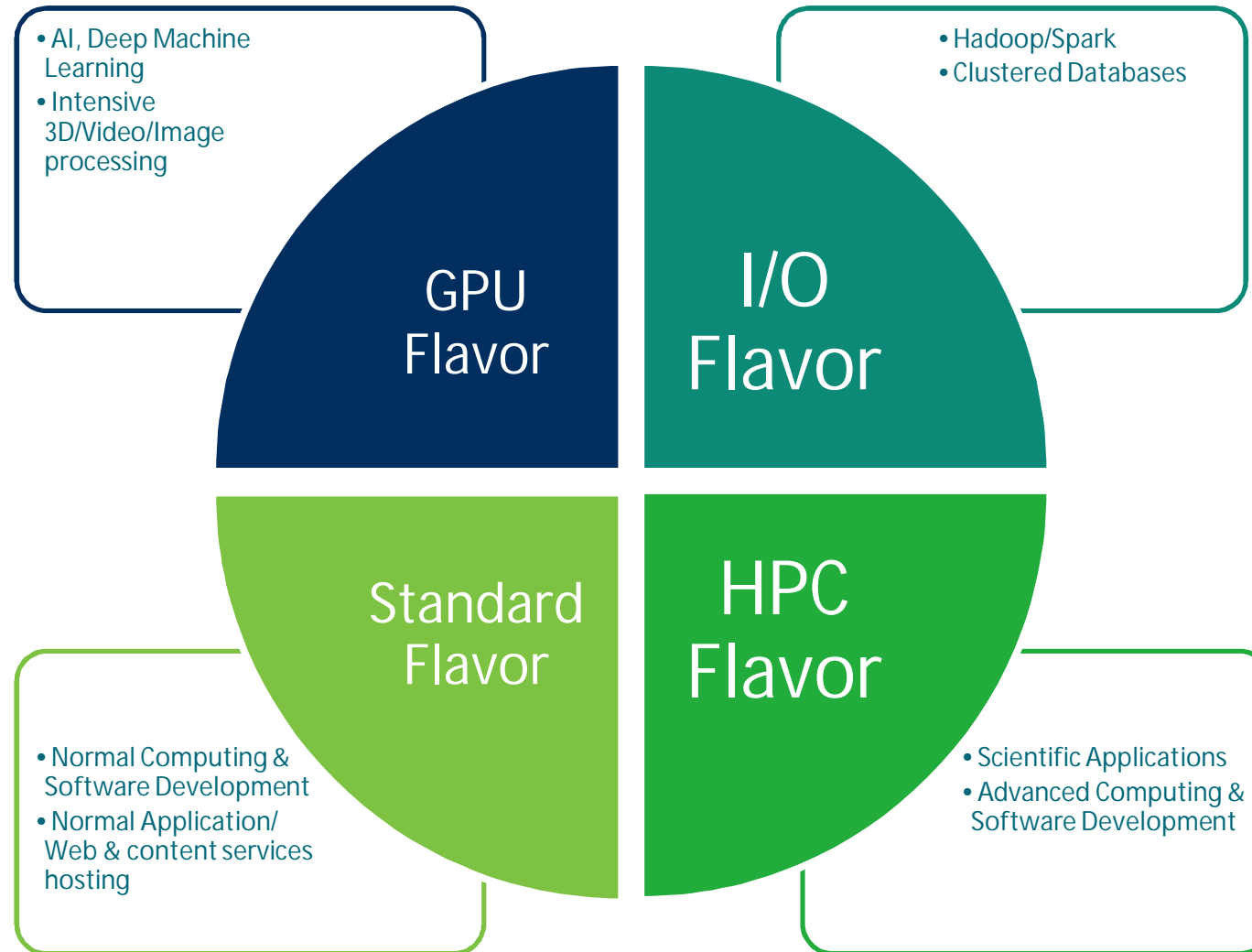
Diverse set of hardware options to support your computing needs

Pouta WebUI

Pouta CLI

Computing services

- Rahti container cloud (https://rahti.csc.fi)
  - Kubernetes/OpenShift environment for running containerised applications
  - Public beta, production in 2020

okd

Search Catalog
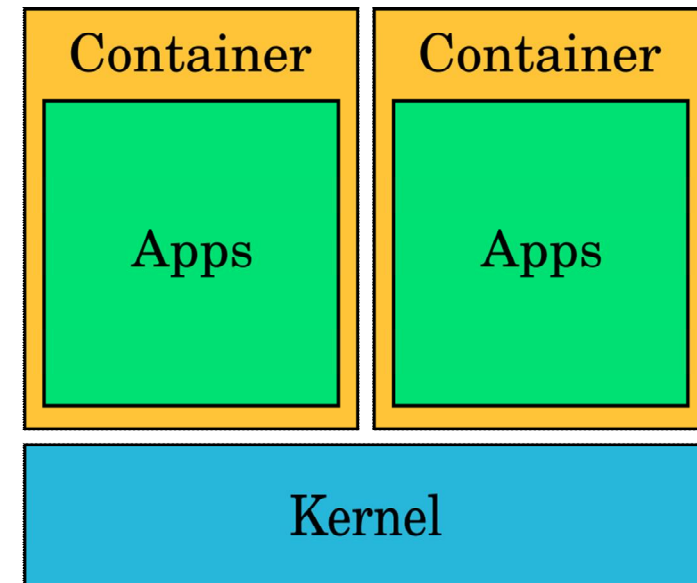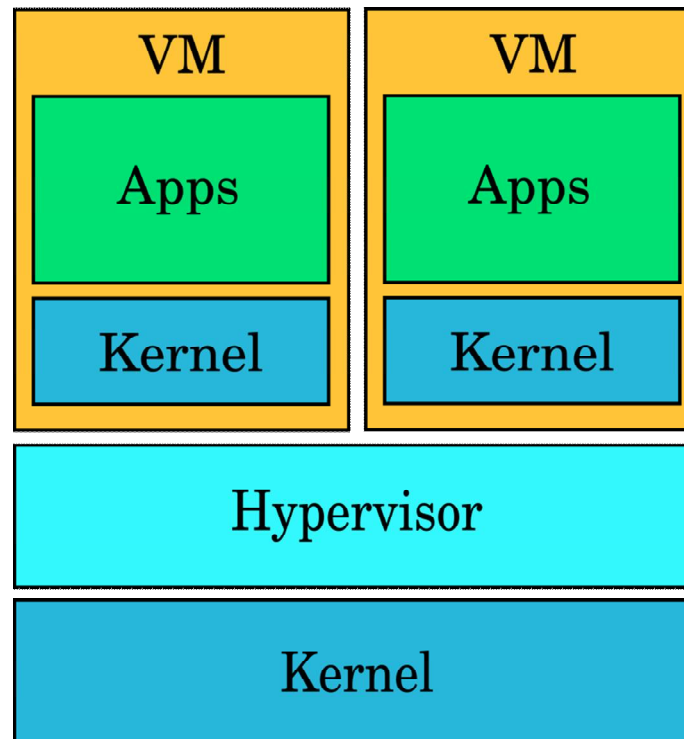
Browse Catalog

All   Languages   Databases   Middleware   CI/CD   Other

Filter ⌄   **43 Items**

.NET .NET Core	**.NET** .NET Core + PostgreSQL (Persistent)	**.NET** .NET Core Example	**.NET** .NET Core Runtime Example	3scale-gateway	amp-apicast-wildcard-router
amp-pvc	Apache HTTP Server	Apache HTTP Server (httpd)	Apache Spark	CakePHP + MySQL	CakePHP + MySQL (Ephemeral)

Why CSC cloud services

Support modern DevOps,CI/CD tools

Support in Finnish, Swedish and English

Services hosted in Finland

High performance Funet Network

Training courses

Serving Cloud computing needs since 2013

Free for universities to do open research