FED4FIRE
Federation for FIRE Plus

Service Oriented Interactive 3D Digital Twin Measuremens (SOI-TWIN)

Fed4Fire Industry 4.0 Experiments

Interactive Cloud Services

Folker Schamel

Spinor GmbH

Experiment Description

Background and Motivation

INTERACTIVE MEDIA

- Instant availability anytime anywhere (no loading times)
- Simple content upgrading
- Live media integration
- Performance
- Device independency

Requires suitable technology strategy: Cloud first instead of lift & shift

MARKETS

- Industry 4.0
  - 3D
  - Natural and visual UIs
- Retail / Configurators in the Configure Price Quote (CPQ) industry:
  - Visual and interactive
- Training, education and entertainment, etc.

Concepts and Objectives

CORE GOALS

Functional and performance analyses of interactive media service deployments in cloud infrastructures

Based on Service Oriented Interactive Media Architectures

SPECIAL ASPECT

- High performance access to GPUs from within WMs running Window.
  (OS choice due to of WMF for hires video, DirectX etc.)
- Selected initial performance measurements of service oriented interactive media deployment architectures

Experiment Setup

SOIM DEPLOYMENT

- Virtual Wall of imec
  - gpunodes n083-01, n083-03 - n083-10 (dependent on availability)
- Multi-user service oriented interactive media architecture
- WM-based
  - Windows Server 2012 WM on Ubuntu 16.04

TESTING & MEASUREMENTS

- Automated packaging, image creation, deployment, startup
- Functional testing of service deployments
- Performance testing of service deployments and startup

Project Results

Failure of Usable GPU Access

LESSON LEARNED: CONTAINERS INTEAD OF VMS

- Our service was running successfully within windows WM hosted by Linux WMs on the Virtual Wall.
- But no satisfactory access to the GPU of the gpunodes.
  - Only basic fall-back drivers, which are not suitable for high-performance interactive rendering and media applications
  - Analysis: This is a general challenge of VMs (not OS specific)
- Conclusions:
  - Tighter integration of base system and application container needed
  - Switch to containers (e.g. Windows containers)

Deployment Performance

RESULTS: IMPORTANCE OF AUTOMATIZATION AND TRAFFIC

- Automatization of packaging and deployment works very well.
  - Python based & shell scripts
  - Virtual Wall and Fed4FIRE connection and tunnel management
    - Automated SSH
  - Abstraction layers: Content, packaging, transfer, startup, logging
  - Essential for rapid workflow and heterogeneous application pool
- Brute-force upload is slower than expected: An hour and more
  - Specific for our use cases. This is what we want to know about.
  - Incremental updates hard to do on image base and breaks isolation when doing on content package base

Startup Time Measurements

RESULTS: DEPLOYMENT OUTSHADOWS STARTUP

- Image upload: 1:20 - 1:30 hours
- Node availability wait times: Typically 12 - 25 minutes
- WM startup times: 1:20 - 1:50 minutes
- Service executable startup time: 12 - 14 seconds
- Service content loading time (base & specific): 2:15 - 2:33 min.
- Shutdown: 35 - 50 seconds (not really relevant in practice)

Conclusion: Sharing data for elasticity of multiple applications

- VMs not optimal: Choice between different images vs. separate content management → Lesson learned: Containers?

Business Impact

Choice of Fed4FIRE

STRATEGICAL ASPECTS

Preference of research oriented partners

- Developing deeper technology for competitive advantage
- Know-how ingest
- Recruiting of talents

Fed4FIRE is the leading platform for such testbeds

ECONOMIC ASPECTS

- Lack of own cloud infrastructure
- GPUs on commercial clouds are expensive and costs are partly unpredictability (for R&D over multiple months)

Commercial Benefit of New Information

FUNCTIONAL

- New knowledge of non-trivial deployments of GPU-based applications on scalable infrastructures

- New knowledge of non-trivial node management and orchestration

PERFORMANCE

- Performance characteristica
  - For our specific use cases and application data
- Basis for use case business analyses and pitches
  - Especially service oriented architectures vs. lift & shift

→ Relevant for commercial battlefield for cloud based interactive media applications

Impact on Technology Strategy

IMPACT ON TECH STRATEGY

Experiments had large impact on future technology strategy:

- Leads to containerization
  - Very tricky and far-reaching consequences due to high-performance Windows GPU access
  - Better for elasticity of multiple applications

IMPACT ON FUTURE R&D

- API and architecture design for multi-layered deployment automatization
- Architecture design for smarter orchestration
  - Combination of own tech, Kubernetes, VPNs
- Leads to new focus on multi-cloud

Exemplary Benefits in Specific Markets

POSITIONING IN AUTOMOTIVE

Specific pitch in automotive

- Addressing specific problems of product configurators
- Solid business arguments based on numbers of Fed4FIRE experiments (SOIM vs lift & shift for interactive rendering apps)

NEW COOPERATION

New research project cooperation with University of Münster

- Would not have been possible without the deployment lessons learned from our experiments on fed4FIRE

Feedback

Node and Image Management

VERY USEFUL. MINOR ISSUES.

- Node management worked very well.
- Image management system worked very well.
- Relevant nodes are often not available
  → "Popular" nodes are used by many researchers
- Issues:
  - Sometimes unclear error message,
    e.g."Node mapping precheck failed!"
  - Launching a Physical Node with Windows (by changing the disk_image urn to a corresponding windows urn) on the n083-01 node failed: Never became ready, waited for 1h.

Networking Architectures

DIRECT NODE ACCESS FOR AUTOMATISATION

- Useful tools for manual access
  - E.g. the nice shell access feature
- But we needed workarounds for full automatization
  - Automated extraction of gateway-ip via ssh trick only.
  - Software seems to be mainly focused on manual access.

    → More focus on automated access would be useful.
    - This might be low hanging fruits: Easy to do, but very beneficial.
- We needed workarounds for public IPv4 access of nodes
  - Direct access would be nice
  - Although we understand that this is not that easy

Hardware & OSes

GREAT LINUX SUPPORT. SAME FOR WINDOWS WOULD BE NICE.

- Yes, Linux is still the main choice for cloud applications.
  But Windows is getting more and more important for interactive media applications.
  - Superior hardware-accelerated media features: WMF, DirectX
- GPUs are getting more and more important.
  → More GPU nodes would be very useful
  - Media application and rendering (as our use cases)
  - Deep learning

Isolation, VMs, containers, orchestration

WISH LIST

- Support for Windows docker containers
  - For many good reasons containers are replacing WMs
  - Windows containers on physical Windows machines
- Kubernetes support
  - Linux and Windows nodes
  - No need to provide Kubernetes. Own installation is most useful.
    - Own choice of specific Kubernetes versions.
    - More control (e.g. custom Kubernetes build for device plugin support on Windows nodes)

FED4FIRE

QUESTIONS ABOUT SOI-TWIN ?

WWW.FED4FIRE.EU

Co-funded by the European Union

Co-funded by the Swiss Confederation

This project has received funding from the European Union’s Horizon 2020 research and innovation programme, which is co-funded by the European Commission and the Swiss State Secretariat for Education, Research and Innovation, under grant agreement No 732638.