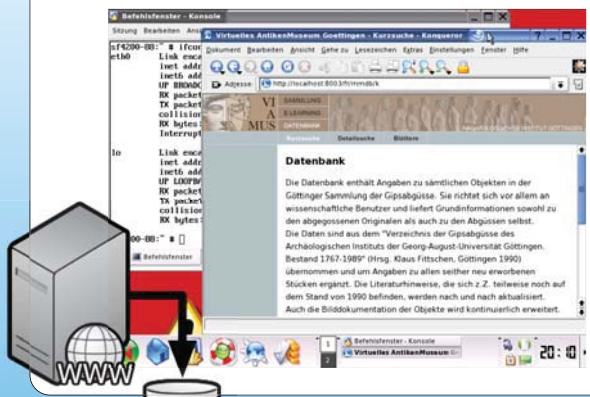
Preservation of Web Content
An Emulation-based Case Study

Dennis Wehrle, Thomas Liebetraut and Klaus Rechert
University of Freiburg, Germany

Original Environment (Server)



Technical Metadata

```
<networkEnvironment
xmlns="http://bwfla.bwl.de/common/datatypes">
<id>9997</id>
<description><title>Viamus Demo</title></description>
<network>

  <!-- client -->
  <emulatorNode>
    <hwaddress>10:23:45:67:89:10</hwaddress>
    <emulationEnvironmentRef>
      <id>9999</id>
    <emulationEnvironmentRef>
  </emulatorNode>

  <!-- server -->
  <emulatorNode>
    <emulationEnvironment>
      <id>9898</id>
      <description>
        <title>Viamus Web Server</title>
      </description>
      <arch>x86_64</arch>
      <drive>
        <data>binding://main_hdd</data>
        ...
      </drive>

      <binding id="main_hdd">
        <url>hdl:11270/...</url>
        <access>cow</access>
      </binding>

      <nic>
        <hwaddress>00:14:4f:0f:73:94</hwaddress>
      </nic>
    </emulationEnvironment>
  </emulatorNode>

</network></networkEnvironment>
```

Workflow

1.) Analysis & Preparation

While the machine is still operational, an assessment of hardware, software and configuration (e.g. operating system, memory, network configuration, ...) is required.

2.) Migrate physical machine to an emulated environment

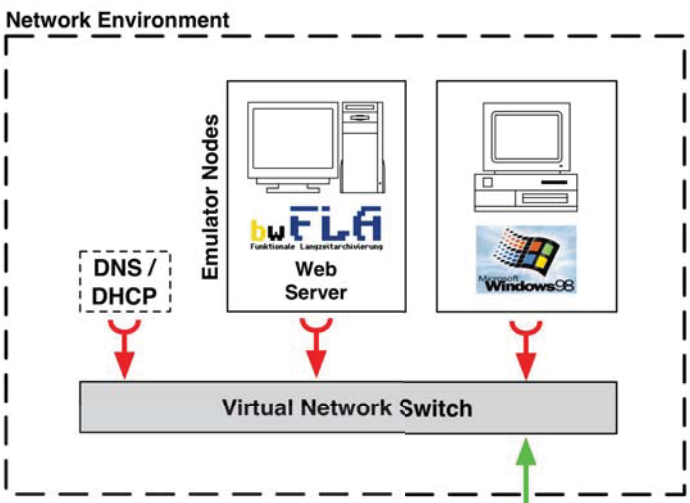
Every system is bound to a specific configuration of its technical environment. Migrating the system into an emulated or virtualized environment is the first step towards a stable environment, as this procedure exchanges the specific hardware configuration with a well documented and well understood configuration (e.g. standard network card).

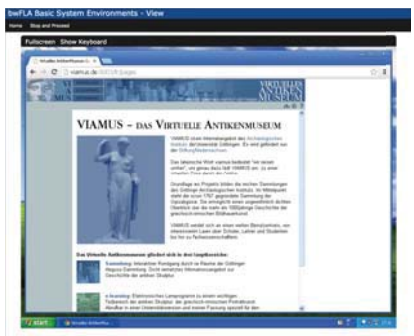
3.) Determine & resolve dependencies

To enable long-term preservation all explicit as well as all implicit dependencies and their functional expectations, i.e. connection to data storage or external data base, have to be identified. Ideally all dependencies should be dealt with to become independent of future changes.

4.) Generate technical metadata

An EaaS instance should be replicable deterministically, i.e. run the same configuration later without configuring all components again. Hence a comprehensive description of an emulated environment is needed. This allows for exchanging EaaS components without losing existing environments and to outlive technical life-cycles.





Access example using emulated client

Re-enact on demand

- complete environment citable (HDL)

- efficent preservation strategy for complex networked environments

-computing costs scale with access demand

- building blocks for a virtual server ecosystem

Instructions

1. Scan QR code
2. Wait until the system has started
3. Start browser and enter: **viamus.de:8003**