

# RELATÓRIO

## Aditivo 01 / E

Arquitetura de hospedagem e  
manutenção do ambiente de  
disponibilização de acervos e  
preservação digital

Maio 2018

---

O presente documento evidencia a realização da segunda etapa do **Projeto Plataforma Acervo: Inventário, Gestão e Difusão do Patrimônio Museológico - Aditivo 01**.

O Projeto é desenvolvido por meio de parceria firmada entre o Instituto Brasileiro de Museus - Ibram e a Universidade Federal de Goiás - UFG, por meio de Termo de Execução Descentralizada - TED.

A descrição técnica da etapa no TED é **“e) Arquitetura de hospedagem e manutenção do ambiente de disponibilização de acervos e preservação digital”**.

---

## SUMÁRIO

1. Introdução.....	4
2. Prospecções.....	6
3. Simulações.....	16
4. Conclusão.....	21
Referencias, containers e virtualizações.....	22

## 1. Introdução

Com o objetivo de apoiar o IBRAM no desenvolvimento de uma estratégia de repositórios em rede para acervos digitalizados da área da cultura, o Ministério da Cultura firmou em Brasília, no dia 17 de outubro de 2016, o Termo de Execução Descentralizada (TED) com a Universidade Federal de Goiás com vistas à realização do Projeto **PLATAFORMA ACERVO: INVENTÁRIO, GESTÃO E DIFUSÃO DO PATRIMÔNIO MUSEOLÓGICO**.

Em um primeiro momento, este TED executou:

a) Análise de requisitos e metodologia participativa para especificação técnica de funcionalidades a serem customizadas no sistema livre Tainacan;

b) Customização e implementação dos requisitos analisados dentro do sistema livre Tainacan+Museu.

O presente relatório tem por finalidade apresentar o resultado da execução do **Produto E do Aditivo 1 deste TED**, cujos processos consistem em:

c) Análise de requisitos e metodologia participativa para especificação técnica das funcionalidades a serem customizadas e integradas ao Tainacan dos instrumentos de coleta sobre os museus brasileiros;

d) Customização e implementação dos requisitos no sistema livre Tainacan e Mapas Culturais;

**e) Arquitetura de hospedagem e manutenção do ambiente de disponibilização de acervos e preservação digital.**

Neste documento será descrito e discutido, portanto, um modelo para infraestrutura de armazenamento de dados do IBRAM. Para a constituição deste modelo foi realizado um levantamento e discussão de um conjunto de tecnologias com suas respectivas características, vantagens, desvantagens e sua aplicabilidade no contexto das necessidades do IBRAM.

Além da realização deste levantamento foram feitas estimativas com o objetivo de ilustrar possíveis cenários de expansão do tamanho do acervo do IBRAM, qual o

---

impacto destes cenários no modelo proposto e quais as medidas e características adequadas ao modelo com o objetivo de atender aos cenários estimados.

Dessa forma este relatório é composto por uma seção relacionada a apresentação das tecnologias levantadas com suas respectivas vantagens e desvantagens em relação ao contexto das demandas da infraestrutura do IBRAM. Outra seção voltada para simulações onde foram simulados cenários considerando diversos tamanhos do acervo do IBRAM. Para subsidiar tais simulações foi considerado o perfil do item com base em uma amostra de um acervo<sup>1</sup> oriundo do Museu Histórico Nacional que está sendo continuamente mantido e revisado em uma instância do software Tainacan<sup>2</sup> hospedado na Universidade Federal de Goiás. Finalizando o relatório são realizadas algumas considerações com base em todas as informações apresentadas nas seções anteriores.

## 2. Prospecções

A arquitetura sugerida é composta por um cluster formado por três serviços essenciais, sendo um serviço de aplicação, um serviço de banco de dados e um serviço de backup. Além dos serviços é necessário prover uma infraestrutura de armazenamento dos dados escalável e dedicada aos serviços descritos acima. Os serviços podem ser replicados pelas máquinas que compõem o cluster e assim garantir sua disponibilidade mesmo diante de eventuais falhas.

O serviço de aplicação é composto em suma por dois serviços sendo uma instância do CMS<sup>1</sup> WordPress e outro uma instância do servidor web Apache<sup>2</sup>. O WordPress é responsável por acomodar os blogs dos museus, com toda estrutura necessária, como plugins, temas e demais recursos necessários. A configuração recomendada para o serviço do Wordpress é uma instalação única (WordPress Multisite) para a todos os - ou maioria dos - blogs. Caso seja necessário, podem ser feitas instalações adicionais exclusivas para museus com necessidades muito específicas. Tal configuração permite uma manutenção e monitoramento mais eficientes dos recursos do serviço como temas, plugins, políticas de segurança e rotinas de backup.

A adoção do WordPress, software base para o Tainacan, se justifica por uma série de características e vantagens, dentre elas o fato de ser open source, gratuito, possuir uma comunidade grande de usuários, grande quantidade de temas e plugins disponíveis para diversos fins. No entanto, esta opção também traz desafios, como a necessidade de manter a sua instância do WordPress e seus respectivos plugins e temas devidamente atualizados para que não haja impactos na sua segurança, tornando-a vulnerável a ataques e invasões. Segue abaixo em relação dos softwares, suas respectivas versões e configurações recomendadas que compõem o serviço de aplicação proposto:

- WordPress:
  - Versão 4.9.4 ou superior;
- Apache HTTP Server;
  - Versão 2.4 ou superior.

---

1 Content Management System

2 <https://httpd.apache.org/>

O serviço de banco de dados é baseado em uma ou mais instâncias do servidor de banco de dados Mysql<sup>3</sup>. A adoção do Mysql justifica-se pelas suas vantagens como ser open source, ter bom desempenho, ser seguro e boa aplicabilidade em diversos contextos de aplicação. A configuração recomendada a nível administrativo para este serviço é que para cada base de dados criada para uma respectiva instalação WordPress seja criado um usuário específico com privilégios restritos à apenas aquela base de dados. É extremamente necessário que seja configurado uma rotina de backup diária de toda a base de dados em um determinado diretório que já esteja incluído no conjunto de diretórios em que o serviço de backup atinge.

Por fim o serviço de backup é uma solução baseada no software Bacula<sup>4</sup>. O serviço deve executar três tipos de tarefas de backup em períodos de agendamento e expiração diferentes. O Bacula permite dentre várias configurações a definição de tarefas (jobs) de backup, a definição de dispositivos de armazenamento (storage), a profundidade do backup (total, incremental e diferencial) e a periodicidade da execução destas tarefas.

Para este serviço a periodicidade sugerida é uma execução diária, durante seis dias da semana, com profundidade diferencial, uma execução semanal e uma mensal, ambas com profundidade total. A profundidade da tarefa de backup se refere ao nível de armazenamento das alterações sofridas nos arquivos e diretórios incluídos nas tarefas de backup. Dessa forma, a profundidade total realiza uma cópia integral dos arquivos e a profundidade diferencial faz a cópia apenas dos arquivos alterados. Outro detalhe relevante para o serviço de backup é a inclusão de todos os diretórios pertinentes ao serviço da arquitetura proposta, como os diretórios da instalação WordPress e os diretórios do serviço de banco de dados, incluindo o diretório onde é realizado o backup diário de suas bases.

Em complemento ao serviço de backup é de suma importância a existência de um planejamento para a infraestrutura de armazenamento que possibilite o aumento do espaço disponível de forma flexível e escalável. Ou seja, é necessário durante a definição de qual será o espaço de armazenamento a ser implantado na infraestrutura que a implantação seja realizada de forma a possibilitar a sua expansão do espaço em disco de forma rápida e sem maiores dificuldades.

---

3 <https://www.mysql.com/>

4 <https://blog.bacula.org/>

Em complemento é importante a implantação de um mecanismo de redundância dos arquivos e dados dos serviços através da habilitação do RAID<sup>5</sup>. O RAID é um mecanismo que permite a fusão de vários discos rígidos em um único volume de armazenamento com o tamanho. Dentre as principais vantagens da utilização do RAID se destacam a melhoria na performance e a possibilidade do espelhamento dos dados armazenados. Considerando que o RAID possui algumas configurações específicas às necessidades de cada sistema de armazenamento no caso deste serviço recomenda-se a configuração do RAID 5.

Analisando as configurações o RAID 5 em suma possibilita a obtenção de um bom desempenho e tolerância a falhas por meio da distribuição dos dados de uma determinada informação que está sendo gravada por todas as unidades componentes do sistema de armazenamento. Tal distribuição consiste em armazenar partes dos dados por todos os discos e junto com cada parte de dados armazenar bits que possibilitem a reconstituição da informação sem a perda de dados. A principal desvantagem desta abordagem é que  $\frac{1}{4}$  do total de espaço de armazenamento disponível será dedicado a realização do RAID e não poderá ser utilizado para outros fins.

Descritos os serviços que compõem a solução sugerida, o próximo aspecto a ser apresentado é a arquitetura que os engloba. Para isso serão abordadas duas arquiteturas baseadas em virtualização com suas respectivas vantagens e desvantagens. A primeira arquitetura a ser abordada é baseada em máquinas virtuais (VMs) e a segunda baseada em contêineres.

A arquitetura baseada em máquinas virtuais possibilita por meio de uma camada de abstração virtualizada a criação de múltiplos ambientes virtuais isolados com seus respectivos recursos, como sistema operacional, memória, espaço em disco, processamento, interfaces de rede e outros periféricos.

Tal camada de abstração é implementada por um hypervisor<sup>6</sup> que executa sobre o sistema operacional instalado no servidor ou máquina hospedeira. Além disso, o hypervisor fornece mecanismos que permitem que cada máquina virtual tenha seu próprio sistema operacional completo, que por sua vez interage com todos os hardwares virtuais disponibilizados pelo modelo de virtualização em nível de máquina. É importante ressaltar que o sistema operacional instalado dentro de uma máquina

---

<sup>5</sup> Redundant Array of Independent Disks

<sup>6</sup> Dentre os exemplos de softwares que atuam como *hypervisor* temos o *Hyper-v*, *Xen*, *Vm Ware* entre outros.

virtual fará interação com os hardwares virtuais e não com o hardware real. A figura abaixo exemplifica como é estruturada uma arquitetura baseada em máquinas virtuais.

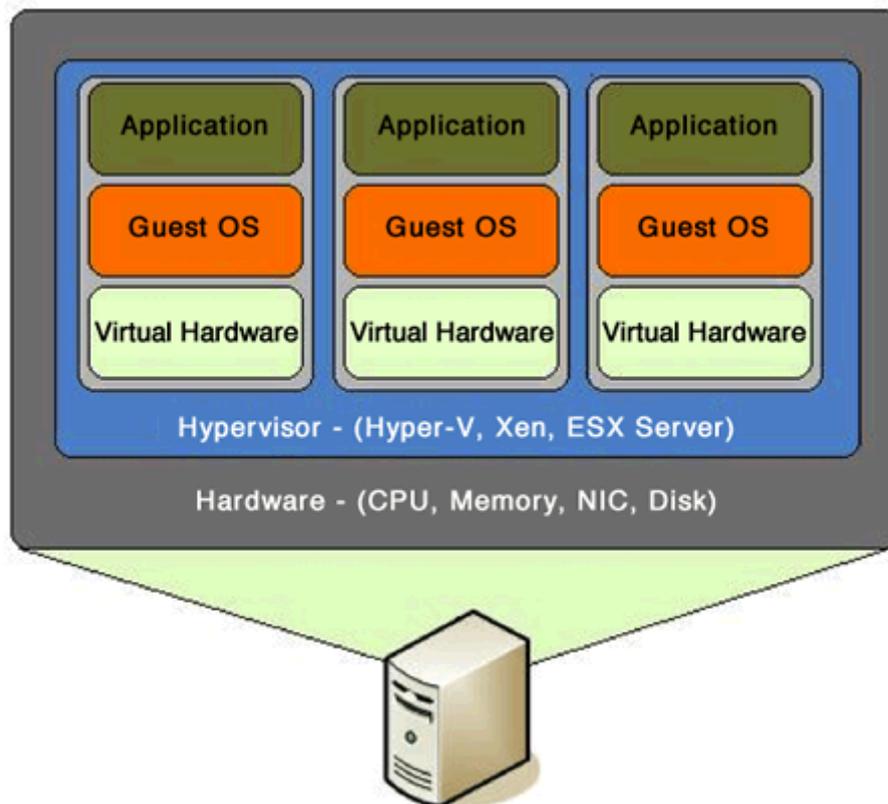


Figura 1: Exemplo da arquitetura baseada em máquinas virtuais.  
Fonte: Gomes (2015)

Dentre as principais vantagens da utilização de máquinas virtuais está o aproveitamento de recursos, o isolamento e a portabilidade. O aproveitamento de recursos caracteriza-se pela possibilidade de execução de várias máquinas virtuais em uma mesma máquina física e especificar quanto recurso físico cada ambiente virtual utilizará e até mesmo aumentar gradualmente em caso de necessidade pontual. O isolamento assegura que as aplicações e serviços que serão executados dentro de uma máquina virtual não poderão interferir no sistema operacional original e nem em outras máquinas virtuais. O que possibilita a construção de um cluster de alta disponibilidade onde devidamente configurado o balanceamento de carga entre as máquinas virtuais torna possível o redirecionamento da execução de um serviço para algum nó do cluster quando alguma máquina virtual fica indisponível. E por fim a

---

portabilidade se resume pela facilidade de migração das máquinas virtuais por meio de mecanismos providos pelo hypervisor.

Dentre as desvantagens apontadas para a utilização das máquinas virtuais podemos destacar a utilização de espaço em disco, consumo de memória RAM e o custo de desempenho. Em relação a utilização do espaço em disco, como cada máquina virtual possui seu próprio sistema operacional e seus respectivos arquivos e diretórios o espaço em disco necessário para alocação de cada máquina virtual pode ser grande. O consumo de memória RAM também é consideravelmente alto levando em conta que para cada máquina virtual uma parte da memória física deverá ser alocada. E por fim, como a arquitetura é baseada em uma camada de virtualização que é implementada entre o sistema operacional e o hardware da máquina física o custo de processamento é superior a uma arquitetura onde não se utiliza da virtualização do ambiente.

A próxima arquitetura a ser detalhada é a baseada em contêineres. Tal arquitetura é composta por uma camada de virtualização em nível de sistema operacional, que em vez de propor a implantação de máquinas virtuais se baseia na utilização de contêineres. Tal modelo tem sua origem datada desde meados da década de 80 e ao longo de todo esse período tem utilizado softwares específicos para implantação dessa camada de virtualização que possibilita a execução dos contêineres, sendo atualmente o software mais utilizado o Docker. A Figura 1 a seguir exemplifica esta arquitetura.

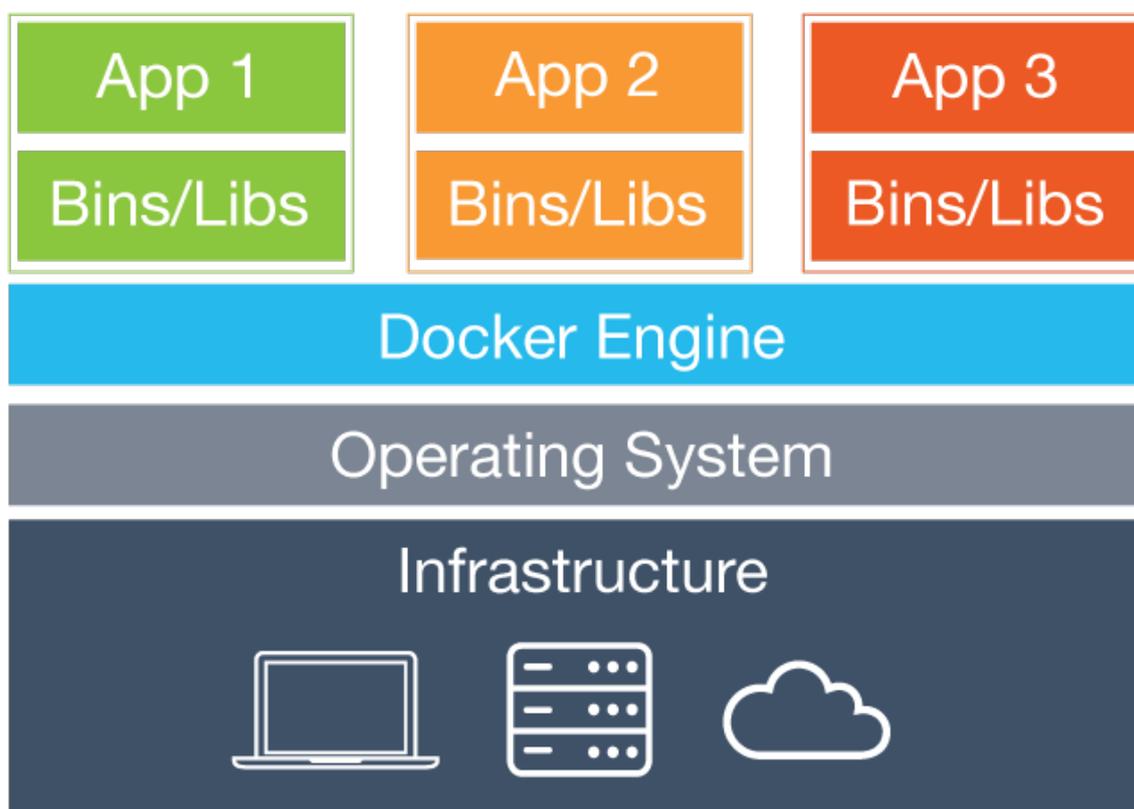


Figura 1: Exemplo da arquitetura baseada em containers.  
Fonte: Gomes (2015)

Nesta arquitetura cada container é apenas um processo em execução em um kernel compartilhado entre todos os outros contêineres. Dessa forma é necessário a implementação de um nível de isolamento entre os contêineres por meio de um mecanismo presente no Kernel denominado *namespaces*<sup>7</sup>, que cria ambientes isolados entre os contêineres, ou seja, os processos em execução em um container não terão acesso aos recursos de outros containers, a não ser que isso seja expressamente liberado na configuração de cada ambiente.

Em complemento, para se evitar a exaustão dos recursos da máquina por apenas um container isolado, utiliza-se também a funcionalidade *cgroups*<sup>8</sup> do kernel, que é responsável por criar limites de uso do hardware disponível. Dessa forma é possível coexistir no mesmo sistema operacional diferentes containers, sem que um afete diretamente o outro por uso exagerado dos recursos compartilhados.

7 <http://man7.org/linux/man-pages/man7/namespaces.7.html>

8 <http://man7.org/linux/man-pages/man7/cgroups.7.html>

Outro detalhe relevante a ser destacado nesta arquitetura é a utilização da plataforma Docker como mecanismo responsável por prover a camada de virtualização que engloba todos os contêineres e gerencia todos os recursos necessários às suas respectivas execuções. O Docker possui um conjunto de mecanismos que objetivam facilitar o desenvolvimento, implantação e execução de aplicações em ambientes isolados agilizando o seu processo de criação, manutenção e modificação do ambiente completo da aplicação.

Dentre tais mecanismos se destaca a possibilidade por parte dos desenvolvedores da aplicação de realizarem uma especificação do ambiente na forma de um determinado arquivo que conterá a descrição dessa solução. Nesta descrição conterão informações sobre a disposição da aplicação neste espaço, como a porta alocada ao serviço, quais volumes externos serão utilizados para a persistência dos seus dados e outras possíveis necessidades.

Por meio desta descrição da aplicação o Docker permite que a mesma seja convertida e empacotada na forma de uma imagem que pode ser facilmente implantada em outro ambiente independentemente do porte da infraestrutura. Além disso, o Docker disponibiliza um serviço em nuvem<sup>9</sup> onde é possível armazenar e disponibilizar as imagens de forma simples e ágil.

Todos esses mecanismos de descrição e criação de imagens da aplicação combinados ao serviço em nuvem para armazenamento trazem uma perspectiva formidável para o IBRAM no que se refere à manutenção, controle e realização da implantação de seus sistemas pela sua rede de museus. Além disso, permite que um museu ou parceiro que já possua uma infraestrutura própria com a plataforma Docker devidamente instalada realize a implantação de algum sistema de forma ágil e prática, bastando baixar a imagem do repositório do IBRAM no serviço em nuvem e realizar sua execução na plataforma do Docker.

Com as duas arquiteturas apresentadas, é possível perceber que o ponto de conflito entre ambas é baixo tornando possível a sua implementação em uma arquitetura híbrida. É possível provisionar uma máquina física com um servidor de máquinas virtuais, onde serão criadas máquinas virtuais hóspedes, que por sua vez terão a plataforma Docker devidamente instalada em cada máquina para a execução e gerenciamento dos contêineres responsáveis pela execução dos serviços descritos

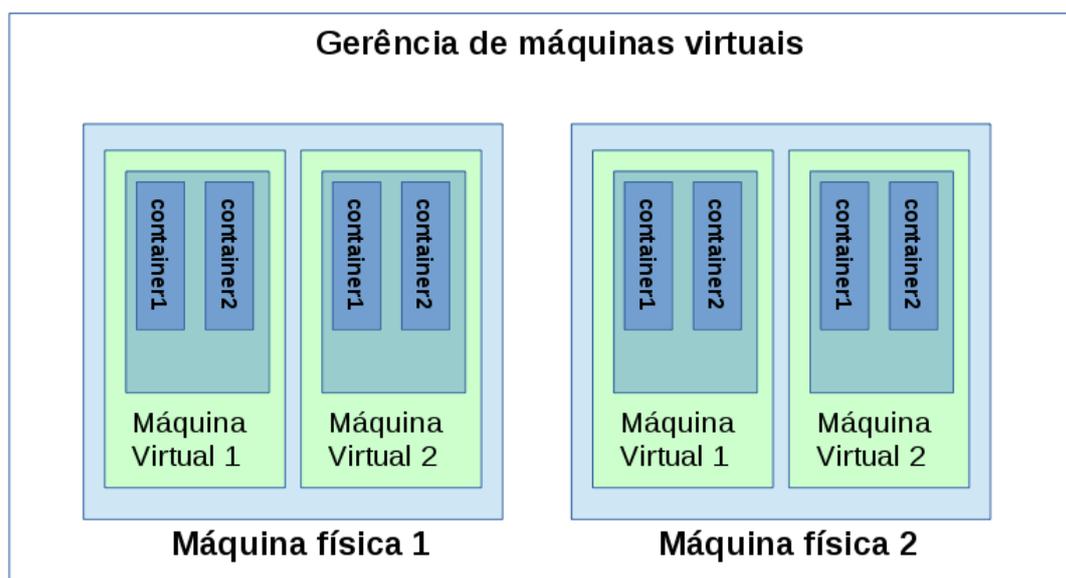
---

<sup>9</sup> <https://hub.docker.com/>

acima. Nesse docker serão disponibilizados os ambientes com seus respectivos serviços segregados, cada um em um container.

Assim é possível implementar vários níveis de isolamento. No primeiro nível é composto pela máquina física que comporta várias máquinas virtuais, ou seja, nessa camada sistemas operacionais interagindo com hardwares virtuais distintos, como placas de rede virtuais, discos, processo e memória. Nesse ambiente teríamos apenas o sistema operacional básico instalado e a plataforma Docker.

No segundo nível de isolamento, a plataforma Docker realiza o gerenciamento dos containers em execução, que por sua vez criam novos ambientes isolados, em nível de processamento, memória, disco e rede. Nesse caso é possível executar numa mesma máquina virtual várias instâncias dos serviços propostos, mas em containers diferentes e isso não seria nenhum problema de boas práticas de gerência de serviços, muito menos de segurança.



Um aspecto relevante neste nível é o gerenciamento de contêineres provido pela plataforma Docker por meio de mecanismos<sup>10</sup> como o modo Swarm<sup>10</sup> e o Compose<sup>11</sup>. O Compose é utilizado principalmente na configuração dos serviços a serem executados na plataforma Docker onde é possível definir para cada serviço suas características

10 <https://docs.docker.com/engine/swarm/>

11 <https://docs.docker.com/compose/>

---

como volume a ser utilizado no armazenamento dos seus arquivos, quantidade de memória, porta alocada e imagem Docker a ser executada entre outros recursos. Além disso é possível por meio do Compose definir uma pilha de serviços a serem iniciados em conjunto e definir as conexões necessárias a execução dos mesmos como no caso dos serviços. Em complemento o mecanismo *Docker service*<sup>12</sup> permite o controle e replicação do serviços alocados na plataforma Docker.

O modo Swarm em suma permite a criação de um cluster de máquinas virtuais ou físicas que estejam executando a plataforma Docker. Entre os principais benefícios do Swarm podemos destacar a possibilidade do balanceamento da execução dos contêineres entre os nós participantes do cluster. Tal balanceamento possibilita a constituição de um ambiente mais tolerante a falhas onde a medida que os serviços são iniciados os seus respectivos contêineres são distribuídos pelas máquinas do cluster.

Finalizada a apresentação dos todos esses aspectos das possíveis arquiteturas para serem implantadas na infraestrutura do IBRAM é possível concluir que uma arquitetura híbrida composta por uma camada de máquinas virtuais e uma camada de contêineres sendo executados nestas máquinas virtuais oferece um ambiente mais seguro e tolerante a falhas permitindo uma maior disponibilidade dos serviços disponibilizados. A utilização da plataforma Docker como ecossistema para a execução e gerenciamento dos contêineres e dos seus respectivos serviços mostrou-se uma ótima opção a ser adotada.

---

12 <https://docs.docker.com/engine/reference/commandline/service/#usage>

### 3. Simulações

Nesta seção objetiva-se a descrição e simulação de cenários onde serão realizadas estimativas do tamanho do acervo do IBRAM. A metodologia adotada para a realização das estimativas em suma se baseia no cálculo médio do tamanho de cada item do acervo e em seguida estimar o tamanho do acervo como um todo. Para tal cálculo serão considerados a quantidade e o tamanho médio de cada anexo que compõe um item do acervo.

Para caracterizar o perfil do item utilizado nestas estimativas e simulações será utilizado como amostra um acervo<sup>13</sup> oriundo do Museu Histórico Nacional que está sendo continuamente mantido e revisado em uma instância do software Tainacan<sup>14</sup> hospedado na Universidade Federal de Goiás. Tal acervo atualmente consta com aproximadamente 22 mil itens dos quais cerca de 510 itens já foram revisados até o presente momento, ou seja, já se encontram com todos os seus respectivos anexos submetidos. De acordo com informações levantadas junto aos responsáveis pela manutenção do acervo, cada item pode possuir de 2 a 4 anexos, na forma de um arquivo do tipo imagem, pode ter um tamanho médio em disco de 5 MB e 6,5 MB. É importante ressaltar que dentro do quantitativo total do acervo existe a possibilidade da existência de itens sem anexos que por possuírem um tamanho inexpressivo não estão sendo considerados nas simulações.

Fechando os atributos considerados nas simulações, também será considerado o tamanho do acervo, tendo como ponto de partida a quantidade já revisada de 510 itens e concluindo com a quantidade total do acervo com 22 mil itens.

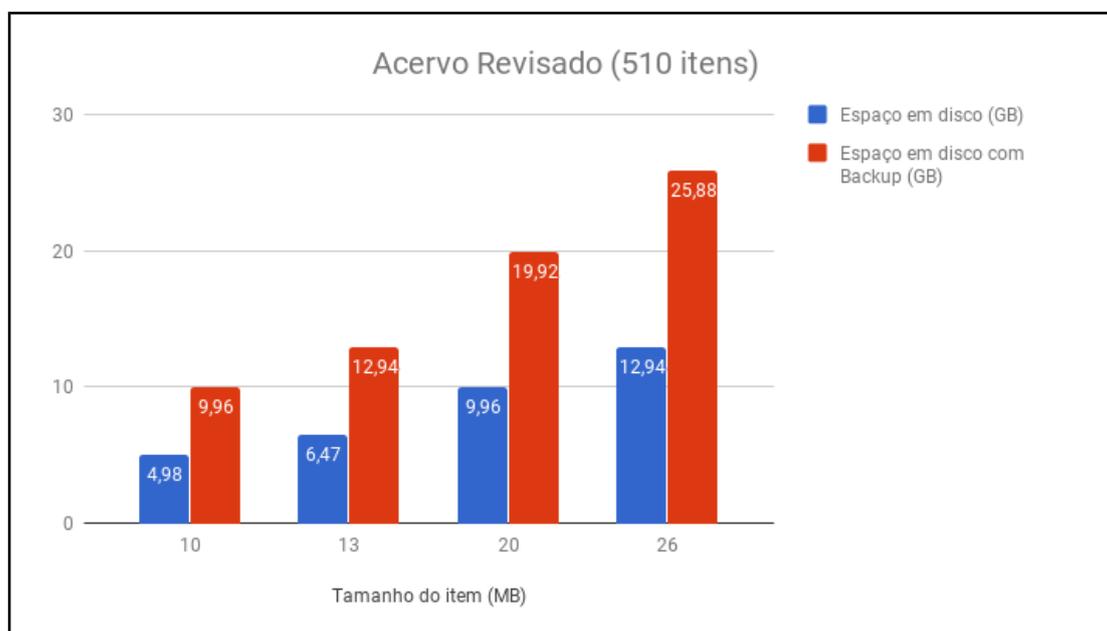
Dessa forma, serão considerados na realização das simulações quatro perfis de item para cada quantidade de acervo. O primeiro perfil é composto por um item com 2 anexos com tamanho em disco de 5 MB e tamanho total do item em disco de 10 MB. O segundo perfil é composto por um item com 4 anexos com tamanho em disco de 5 MB, tamanho total do item em disco de 20 MB. O terceiro perfil é composto por um item com 2 anexos com tamanho em disco de 6,5 MB, tamanho total do item em disco de 13 MB. O quarto perfil é composto por um item com 4 anexos com tamanho em disco de 6,5 MB e tamanho total do item em disco de 26 MB.

---

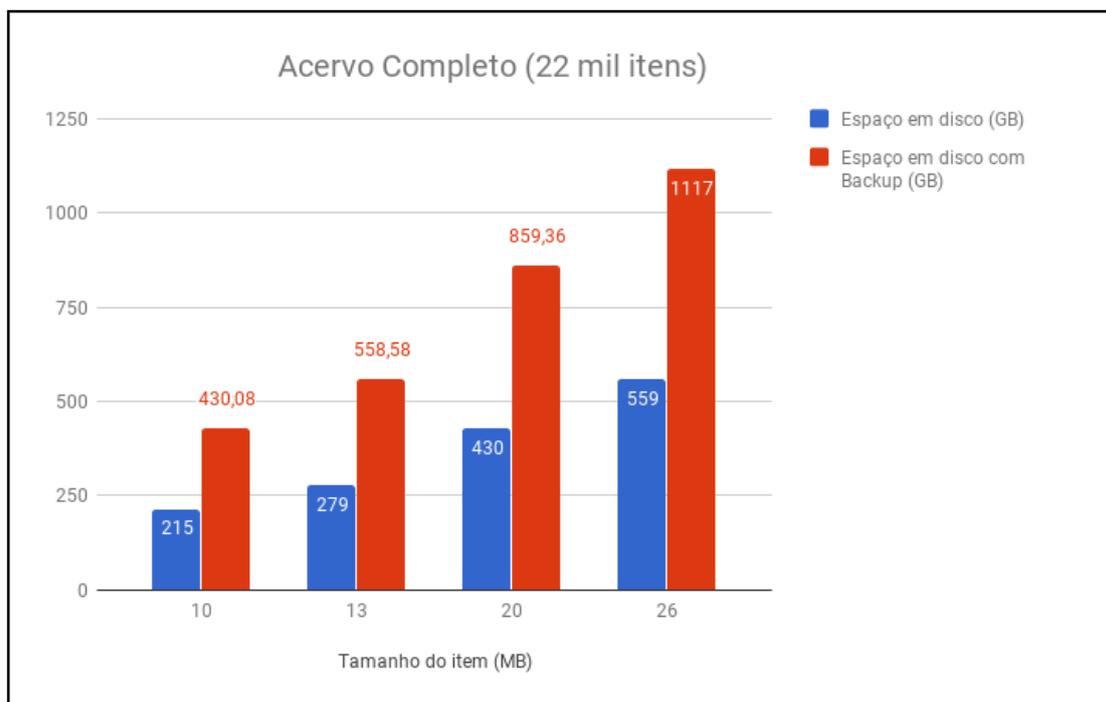
13 <https://mhn.medialab.ufg.br/colecao/itens-seret/>

14 <http://tainacan.org/>

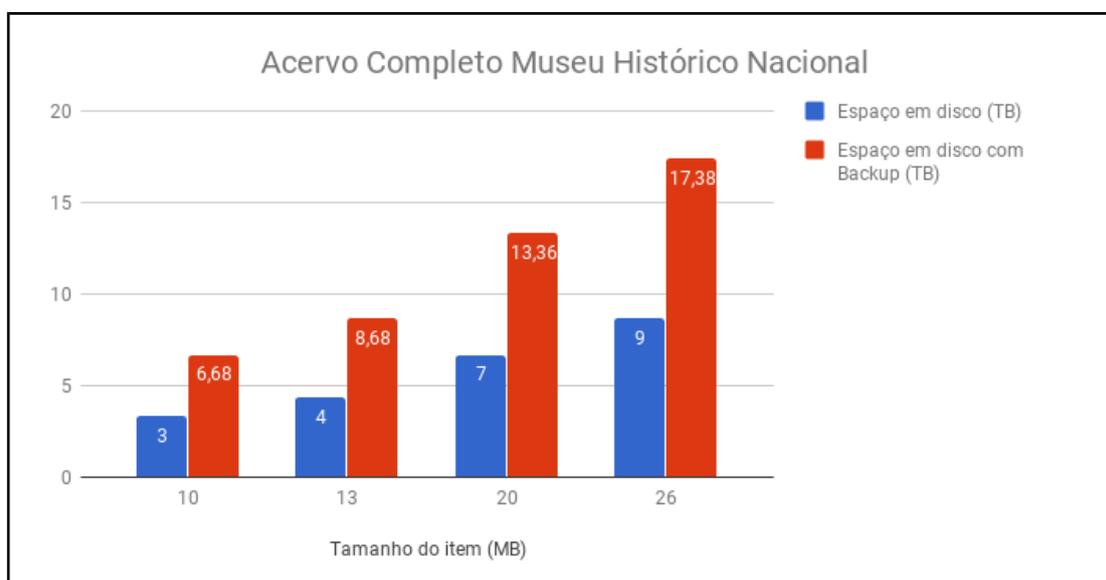
Em todas as simulações considera-se a implantação de uma rotina de backup que realize o espalhamento dos arquivos do item do acervo de forma que se tenha um ciclo mensal com uma cópia completa dos arquivos no início do ciclo e com uma rotina diária de incrementação das alterações sofridas. O gráfico abaixo mostra a simulação realizada com o acervo já revisado que conta com cerca de 510 itens. No eixo horizontal o gráfico possui quatro pares de colunas com estimativas do tamanho do acervo de acordo com o tamanho do item em megabytes (MB). Em cada par das colunas existe uma estimativa de todo tamanho do acervo em Gigabytes (GB) considerando um espelhamento dos arquivos dos itens, ou seja, uma cópia completa de todos os arquivos. Nesta primeira estimativa já se observa um aumento óbvio e gradativo do tamanho do acervo a medida que o tamanho do item também aumenta.



O próximo gráfico mostra a simulação realizada considerando todo o acervo já submetido que conta com cerca de 22 mil itens. Em destaque se observa a última coluna do item com tamanho do item em 26 MB, onde o tamanho do acervo já entra na dimensão dos terabytes com aproximadamente 1,09 TB.



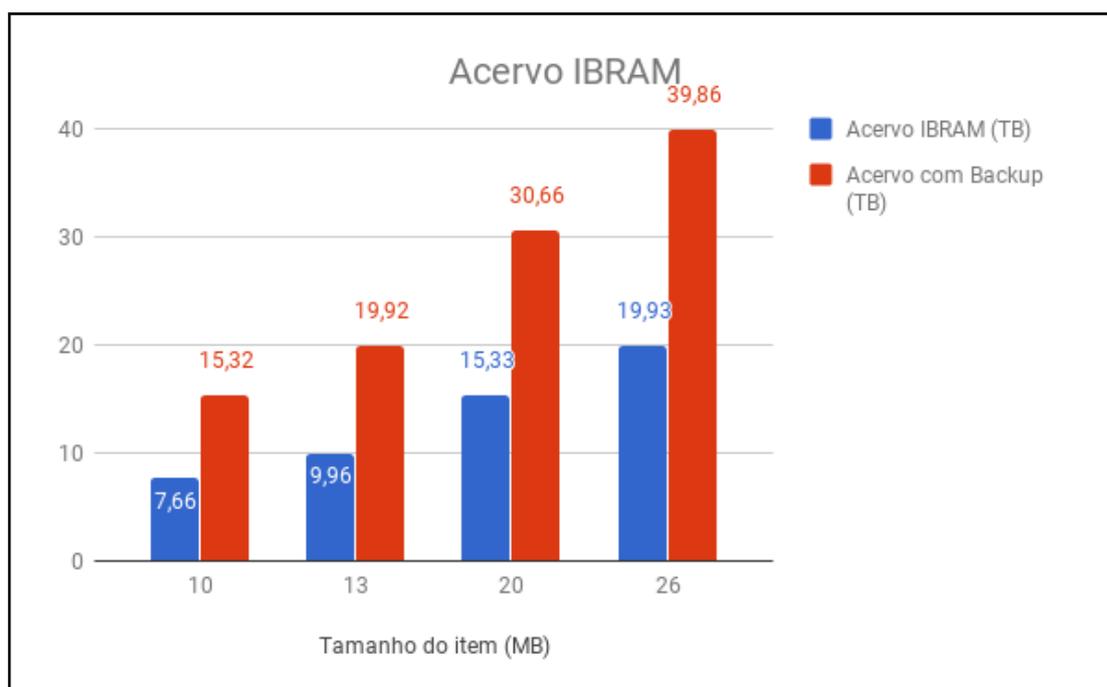
Agora considerando todo o acervo do Museu Histórico Nacional que é segundo seus mantenedores estimado em mais de 350 mil itens o cenário se demonstra segundo o gráfico a seguir.



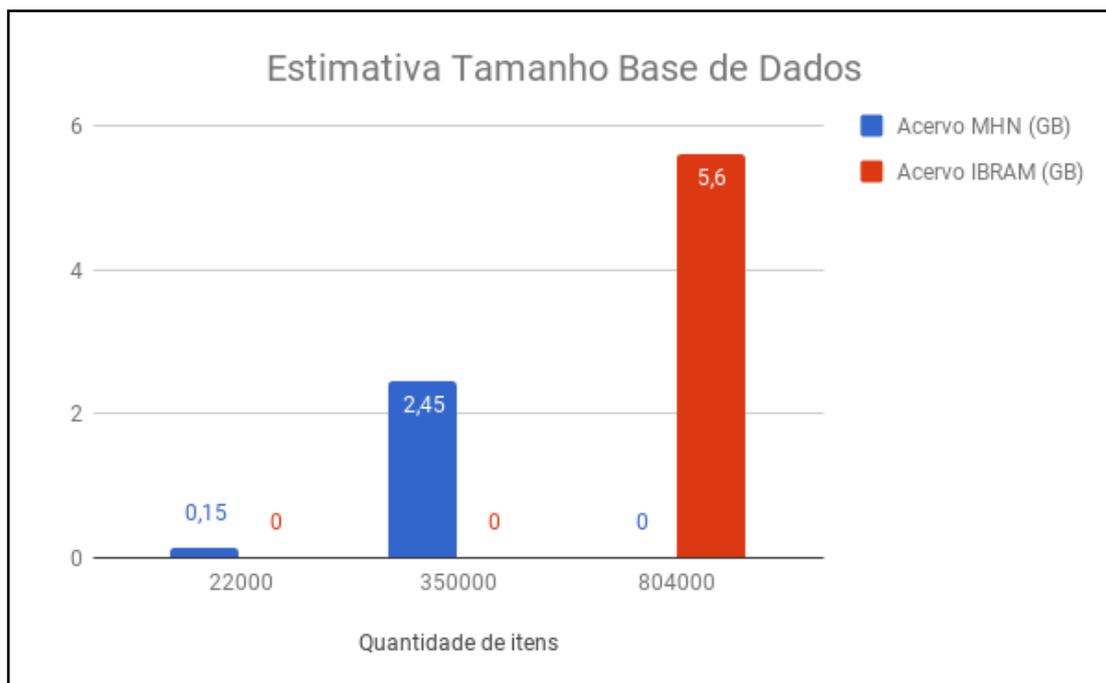
Considerando os cenários simulados acima se observa um crescimento exponencial do tamanho do acervo do MHN a medida que o tamanho do item aumenta o que denota uma tendência esperada. Tal tendência evidencia a importância da implantação de uma solução de armazenamento dos arquivos escalável e com expansão flexível. A tomada do acervo do MHN como referência se demonstrou interessante ao ponto que representa o recorte de um acervo real e permite realizar estimativas com um o maior nível de conexão possível com a realidade dos acervos dos museus.

O próximo cenário a ser simulado considera todo o acervo do IBRAM onde segundo informações levantadas junto aos responsáveis possui uma quantidade aproximada de 804 mil itens. O gráfico abaixo demonstra o cenário simulado.

Uma ressalva importante em relação aos perfis de tamanho de item adotados como referência para as simulações relataram como anexos imagens, neste caso considera-se a possibilidade do item ser composto por outros tipos de arquivos como por exemplo áudios, ou seja, a informação relevante para a estimativa é o tamanho total do item e não a sua composição.



Outra perspectiva a ser considerada nas simulações é a estimativa da evolução do tamanho da base de dados em relação tamanho do acervo. Para a realização desta estimativa foi realizado um cálculo aproximado do tamanho de um item em relação ao tamanho da base e chegou-se ao valor aproximado de 7,36 KB (Kilobytes). Tomando como base o valor do item na base de dados o cenário a ser simulado considera três quantitativos de itens nos acervos, sendo dois relacionados ao acervo do MHN e um considerando todo o acervo do IBRAM sendo tais quantitativos, respectivamente, 22 mil, 350 mil e 804 mil itens. O gráfico abaixo denota esta simulação.



---

## 4. Conclusão

Neste documento procuramos investigar e dimensionar os requisitos técnicos para infraestrutura de hospedagem dos acervos digitais dos museus ligados ao IBRAM.

Do ponto de vista de arquitetura de aplicações, foram apresentadas algumas alternativas e sugerida uma solução híbrida entre máquinas virtuais e containers utilizando a tecnologia Docker, para que seja fácil de escalonar a estrutura e distribuir as aplicações em diversos servidores.

Do ponto de vista da armazenagem de dados, foi realizado um dimensionamento do tamanho necessário para hospedar todo o acervo de todos os museus, de acordo com as informações que temos em mãos sobre as características desses acervos.

Do ponto de vista da armazenagem do banco de dados, avaliamos que é algo com baixo impacto. Mesmo na simulação com 804 mil itens, o tamanho da base de dados chega a apenas 5,6GB nas projeções realizadas. Ou seja, em comparação com a armazenagem de arquivos, esse tamanho é praticamente irrisório. Considerando que pode haver grande variação do tamanho por item, dependendo da natureza do acervo, e considerando que o impacto de infraestrutura de armazenamento é pequeno, sugerimos extrapolar essa projeção e garantir um espaço de armazenamento de pelo menos 20 vezes esse valor para a base de dados.

Já em relação a armazenagem de arquivos digitais, como fotos, documentos, áudios e outras mídias, a projeção mostrou que a demanda de infraestrutura pode ser bastante alta.

---

## Referencias, containers e virtualizações

1. [http://www.mariolb.com.br/blog/\\_static/TCC/TCC-ClusterAltaDisponibilidadeEmLinux-Diego-Vinicius-2012.pdf](http://www.mariolb.com.br/blog/_static/TCC/TCC-ClusterAltaDisponibilidadeEmLinux-Diego-Vinicius-2012.pdf)
2. <https://www.howtoforge.com/tutorial/install-and-configure-drdb-for-network-file-system-replication-debian-8/3/>
3. <https://imasters.com.br/desenvolvimento/vms-vs-containers-quais-diferencas-e-usos/?trace=1519021197&source=single>
4. <http://blogbrasil.comstor.com/o-que-e-conteinerizacao-de-aplicacoes-e-como-isso-influencia-a-ciberseguranca>
5. GOMES, Rafael. CONTAINER VS MÁQUINA VIRTUAL? 2015. Disponível em: <<http://techfree.com.br/2015/11/container-vs-maquina-virtual/>>. Acesso em: 20 mar. 2018. Gomes (2015)
6. TECNOLOGIA LIVRE. Entendendo o docker parte 1. Disponível em: <<http://techfree.com.br/2015/10/entendendo-o-docker-parte-1/>>. Acesso em: 16 mar. 2018.
7. <http://www.develsistemas.com.br/virtualizacao-de-servidores-vantagens-e-desvantagens/>