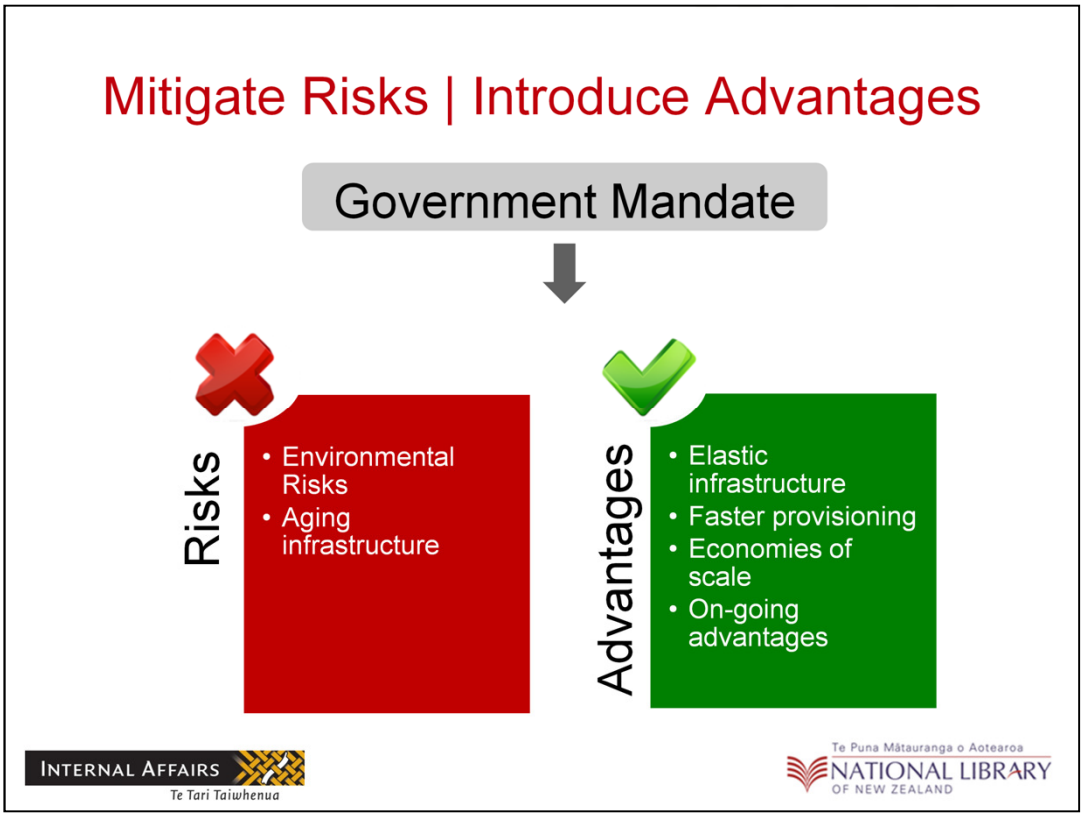
Storage Architecture and Challenges at the National Library of New Zealand

Library of Congress Designing Storage Architectures Meeting
September 22, 2014
Washington D.C., U.S.A.

Cynthia Wu

INTERNAL AFFAIRS
Te Tari Taiwhenua

Te Puna Mātauranga o Aotearoa
NATIONAL LIBRARY OF NEW ZEALAND



In August 2011, the New Zealand government decided to adopt Infrastructure-as-a-Service (IaaS).

The National Library had to begin consider the implications of moving our digital collections to an external data center.

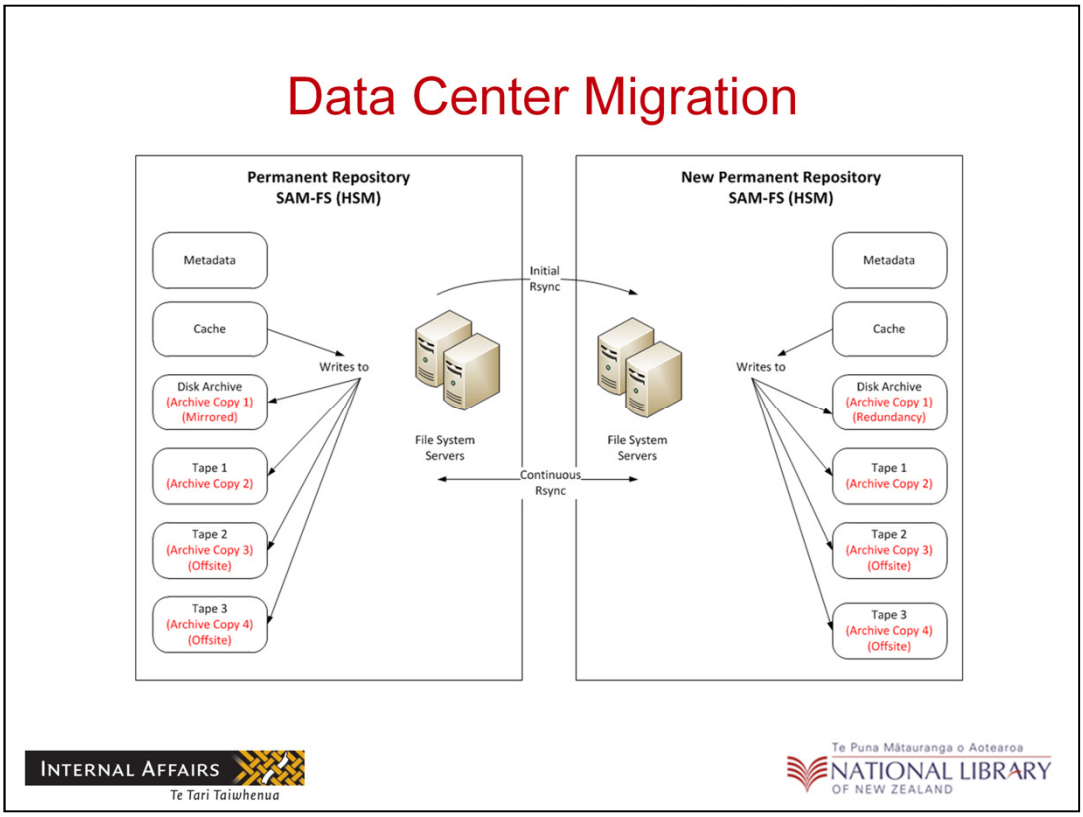
The IaaS initiative are reducing risks and introducing some significant advantages:

Risks:

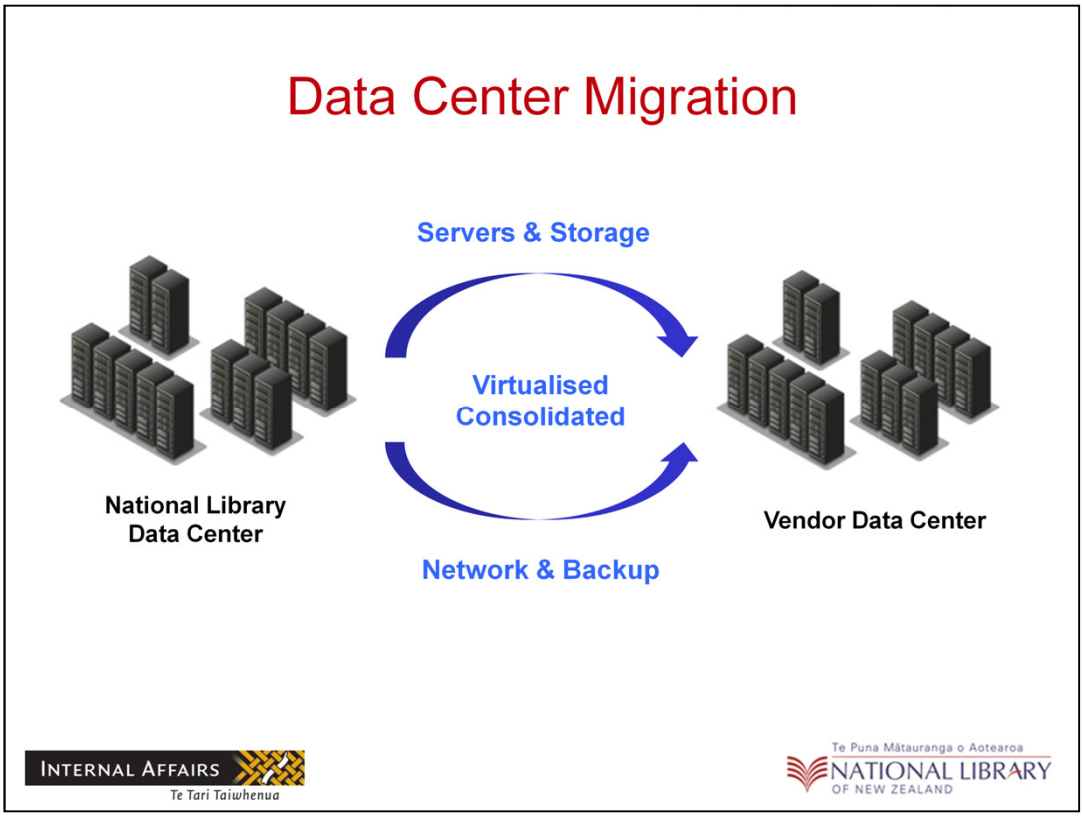
1) Environmental risks - Data centre sits under sewage pipes, earthquakes
2) Aging infrastructure, some close to end of life.

Advantages:

1) The elastic and faster provisioning speed for storage and other infrastructure translates to easier capacity planning and development resource management for the Library
2) The Library can leverage All-of-Government discounts as storage utilization increases.
3) Through Economies of scale, we can even explore future storage innovations with storage provider that previously may have not been possible.



- Pilot project – moving the permanent repositories of our digital preservation systems from our data center to the vendor's data center
- No architectural changes (maintain use of SAM-FS, a Hierarchical Storage Management Filesystem)
- Pilot completed successfully



- Full data center migration of all infrastructure including storage.
- Virtualised servers sharing data via SAN storage according to different tiers that reflect availability and performance requirements.

What happened??

- Lost of Knowledge
- Lack of time to analyze, rationalize, and consolidate
- Right storage tiers for data's purposes
- Backup costs
- Balancing storage cost with usage and performance

INTERNAL AFFAIRS
Te Tari Taiwhenua

Te Puna Mātauranga o Aotearoa
NATIONAL LIBRARY
OF NEW ZEALAND

Data Center migration issues:

1) Lost of knowledge -

- Aging infrastructure hadn't been reviewed for many years
- IT staff turnover throughout the years resulting in lost of knowledge.
- Big impact on resulting architecture that would be transferred to our provider.

2) Lack of analysis, rationalization, and consolidation

- 6 weeks project time
- Lack of time meant no rationalization ro consolidation – data migrated as is, instead of architecting a solution in a structure that would yield improved performance and efficiency.

3) Storage tiers fit for data's purposes

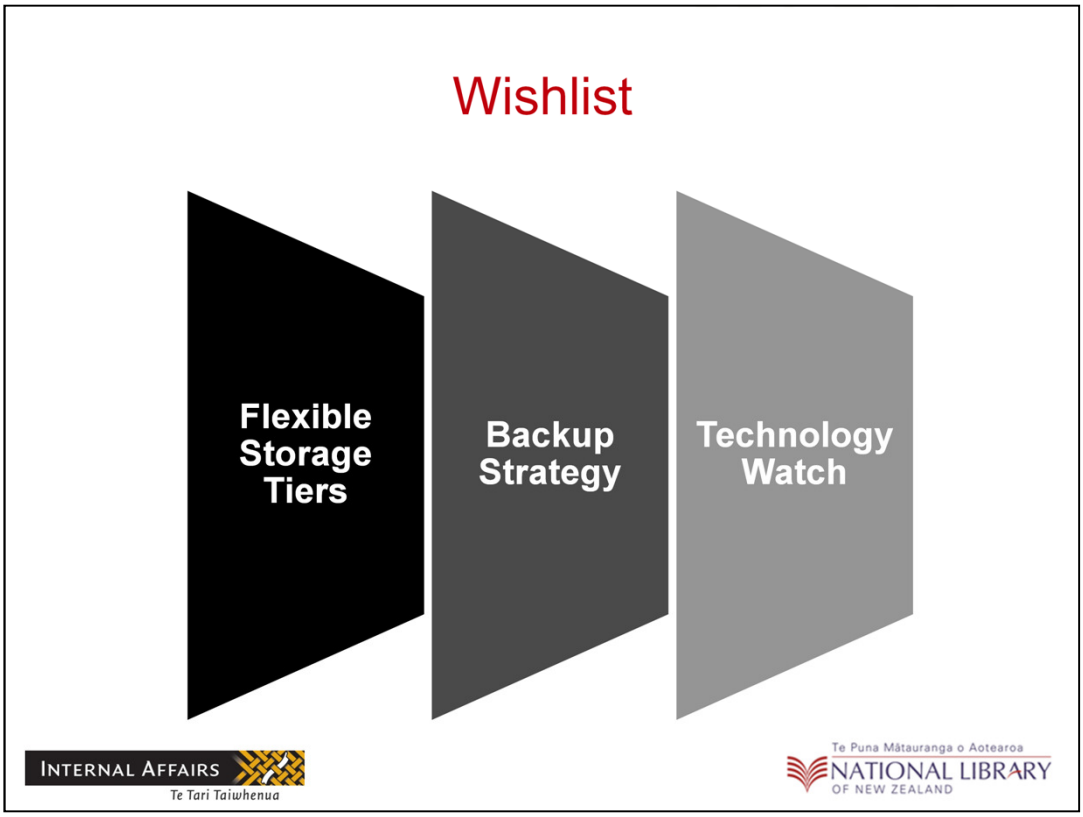
- Severe time constraints led to lack of testing
- Data stored at tiers based loosely on perceived access performance requirements and potential matching tiers at vendor's data center

4) Backup costs

- Significant backup costs compare to actual data storage cost
- Lack of rationalization meant higher backup costs
- Changing organization of data based on backup strategy that will lead to lower backup costs

5) Storage cost versus usage and performance

- Library grow its data continually (no data removal)
- Delivery of material directly from our repositories – not dark archiving
- Challenge of trying to balance the cost of storage versus delivery response time performance.



Flexible Storage Tiers

- Have our data flowing up and down in different storage tiers depending on access demands.
- Store the data at a lower tier and have it flow upward to high tiers when required.
- Pay for data storage at a lower cost for most of the year, and the higher price only for the duration when required.
- Example scenarios: fixity checking and large scale content analysis.

Back up Strategy

- Considering using data replication onto lower tier disks at another geographically separated location, rather than the traditional backup on tapes.
- May allow quicker service recovery time at a lower cost even with possible performance impact of providing services out of the lower tier disks.

Community driven technology watch

- Enables storage users / user organizations to engage in conversations with storage providers.
- Engage in some forward-thinking, planning and a chance to implement potential solutions