Data Capital in eResearch

Maximise the value of your Data Assets

Charles Sevior
CTO, Dell EMC Unstructured Data Solutions

Eric Burgueño
Systems Engineer, Plant and Food Research
A Great Place to live and work!

Novel Ideas
People. Powered. Research.
Substantial investment & opportunity

The result (and often the source) is essentially DATA
Digital Transformation Is Disrupting Every Industry

50%+ of global GDP will be digitized by 2021

80%+ of the world’s data is Unstructured

80% of the world’s data is unreadable by machines
Transforming Data Into Value

- Unlocking data trapped in silos
- Keeping up with data growth
- Creating business impact from data
What is Data Capital?

Data is rapidly becoming organisations’ most valuable asset.
80% of Data is Unstructured

3 DATA TYPES THAT NEED TO BE TREATED DIFFERENTLY

FILE

OBJECT

STREAM
Unstructured Data Solutions

UNIFIED DATA LAKE

SIMPLICITY AT SCALE

EXTRACT VALUE FROM DATA

Dell EMC

ISILON

FILE

STREAM

PROJECT "NAUTILUS"

ECS

OBJECT

Data Management

Extract Value from Data
Plant & Food Research is a New Zealand-based company providing research and innovation to ensure the sustainable growth of the plant and marine-based food industries.

Our Core Purpose: Enhance the value and productivity of New Zealand’s horticultural, arable, seafood and food and beverage industries to contribute to economic growth and the environmental and social prosperity of New Zealand.
By the numbers

Incorporated December 2008
» Based in New Zealand
» Government-owned Crown Research Institute
» Revenue **NZ$146 million** (2016/17)
» Legacy extends back to 1926

More than **900 employees**
» 660+ research staff
» 14 sites in New Zealand
» Representatives in the USA, Australia

Income
» 57% private contracts and royalties
» 43% New Zealand Government contracts
Horticulture helps to build New Zealand’s profile in many overseas markets.
Our science combines market insight with a fundamental understanding of the biology of our food resources.

- **New Cultivar Innovation**
  - Better cultivars faster™

- **Bioprotection**
  - Residue-free pest and disease control

- **Sustainable Production**
  - Sustainable and profitable production systems

- **Food Innovation**
  - Premium foods and beverages

- **Seafood Technologies**
  - Sustainable, premium seafood and marine products
How do we do it?

The Problem

- Sequencing Data exploding
- Traditional SAN Storage not fit for purpose
The changeover

The Analysis
- Trialled Isilon onsite
- Tested competitor products in Sydney
- Reference Site in Perth

The Solution
- Isilon X400 in Primary Data Centre
- Isilon NL400 in Secondary Data Centre
- 10Gb storage network with Arista Switches
- Happy Customers
- Added Corporate File Serving
- powerPlant
powerPlant
## powerPlant

<table>
<thead>
<tr>
<th>Hardware</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU Threads</td>
<td>784</td>
</tr>
<tr>
<td>Total Memory</td>
<td>9 TB</td>
</tr>
<tr>
<td>Total Storage</td>
<td>2.4 PB</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Usage</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Jobs ran</td>
<td>Over 2.6 million</td>
</tr>
<tr>
<td>CPU time (years)</td>
<td>372+</td>
</tr>
<tr>
<td>Storage consumed</td>
<td>900+ TB</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>People</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Core Operations</td>
<td>3</td>
</tr>
<tr>
<td>Application Support</td>
<td>8</td>
</tr>
<tr>
<td>Total Users</td>
<td>240</td>
</tr>
</tbody>
</table>
To-Do List

- Cloud Bursting
- Dealing with inactive data
- Streamline external collaboration
Dell EMC Isilon + ECS + ClarityNow v3

Bring IT and Research together to unlock data capital

- Single view across file and object storage
- On-demand movement from file to object and back
- Proven at scale and speed
Unstructured Data Solutions

Data Management

UNIFIED DATA LAKE

SIMPLICITY AT SCALE

EXTRACT VALUE FROM DATA

ISILON

FILE

STREAM

PROJECT “NAUTILUS”

ECS

OBJECT

ECS

Dell EMC

Unstructured Data Solutions
Dell EMC Isilon

- Single File System & One Namespace
- Global Availability & Protection
- Unmatched Storage Efficiency
- Simplicity & Ease of Use
- Linear Scalability
- Automated Policy Based Tiering
- Unmatched Storage Efficiency
Dell EMC ECS

Unlimited Object Storage

Enterprise Grade Security and Compliance

48% Lower TCO than Public Cloud

Comprehensive API-based Data Access

Native Multi-tenant Architecture

Built-in Metadata Search
Project “Nautilus”

Auto-Scaling of data streams

Improved data durability

Transaction Support

Unlimited Retention

High throughput Write Efficiency

Event handling to eliminate errors
AI / ML / DL Solution Options

- **Ready Solution for AI:**
  - Deep Learning with NVIDIA

- **Augment existing Isilon with F800 All-Flash**

- **GPU-Included**
  - Extreme Deep Learning with NVIDIA DGX-1 & Isilon

**Optional GPU Acceleration (NVIDIA)**

- Compute & Converged
- Hyper-Converged w/ NW
- Existing Servers
- Hyper-Converged

- Massively Parallel IO