



Why do you want so much data and all of it in the middle?

Matthew Rapier

Vice President - CTO



SECURITY & INFORMATION SYSTEMS DIVISION

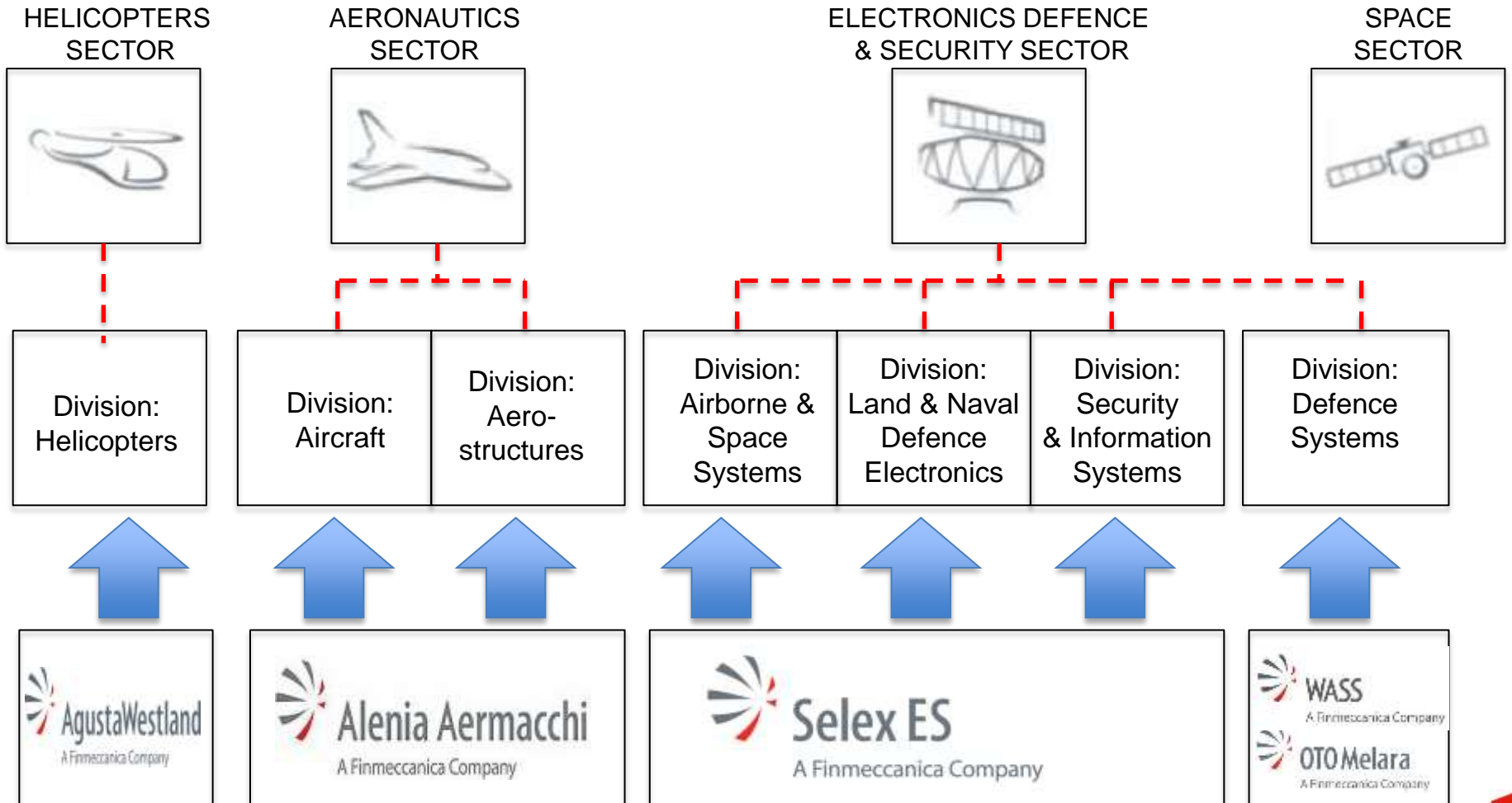
All views presented are those of the author and do not necessarily represent the views of

Selex ES Ltd trading as Finmeccanica



One Company

Finmeccanica's new organisational structure based on the divisional model "One Company". Is effective from 1 January 2016 and is organised into four Sectors and seven Divisions, with a new Governance aimed at centralizing the Group's guidelines and control systems whilst decentralizing the management of the business to the Divisions.





Introduction

- Cyber – at an interesting point in its Hype Cycle
- Time to learn from experience

- Raising a question

- Perspective
 - A Mathematician a very long time ago
 - Over 30 years in ICT



IT in my lifetime



Acorn Atom
1MHz 8 bit
1KB RAM
Cassette



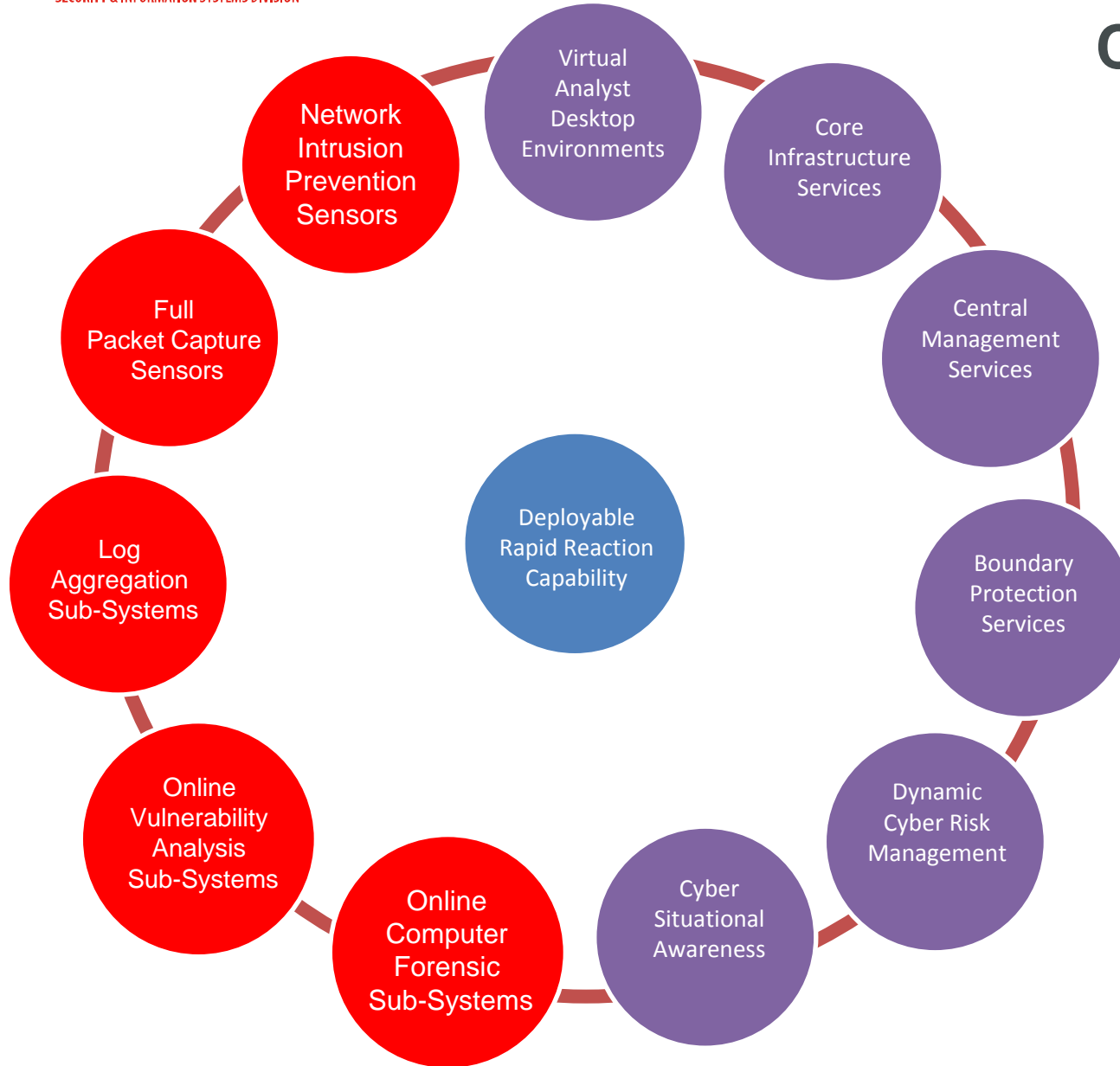
Cray Y/MP 8/432
4 * 200 MHz 64bit
256MB RAM
20GB Disk

Cisco UCS
Flexpod
As much as
you like





Cyber for a major customer



Capability Deployed Across Federated Network

Capability Deployed at Central Sites

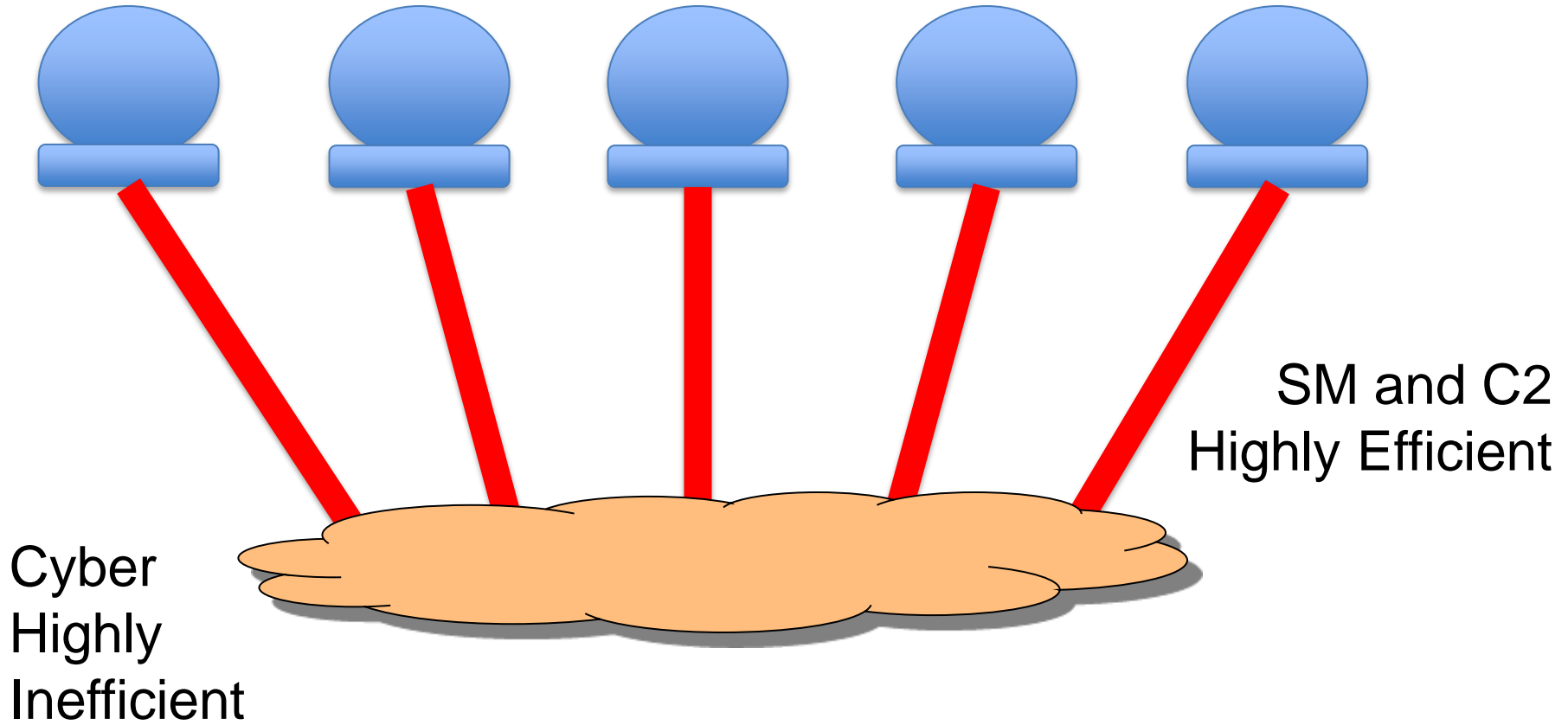


The Analytical Model



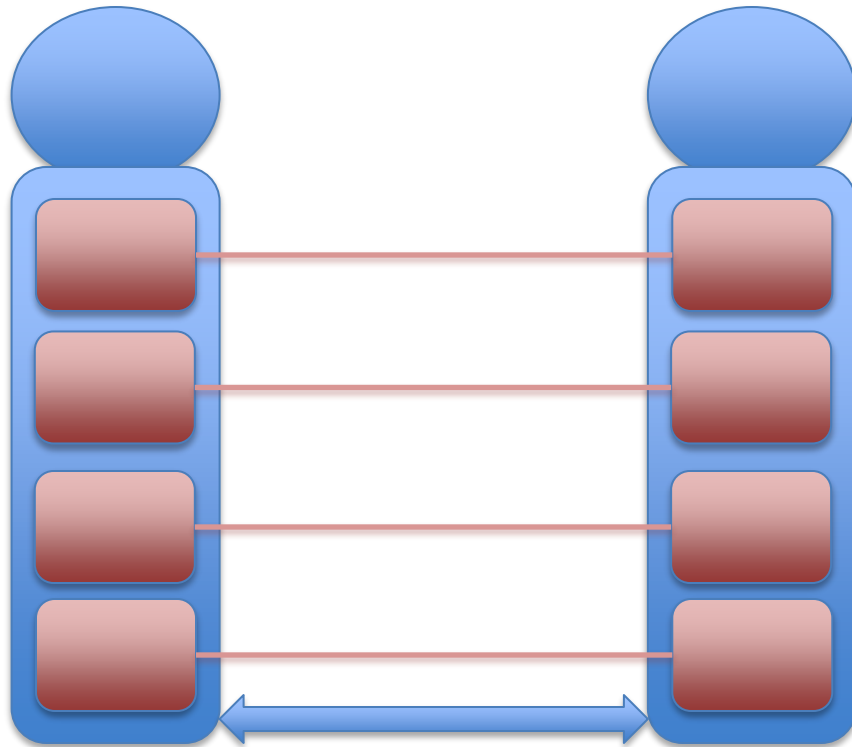


Distributed Architecture





Effective Distributed Architecture

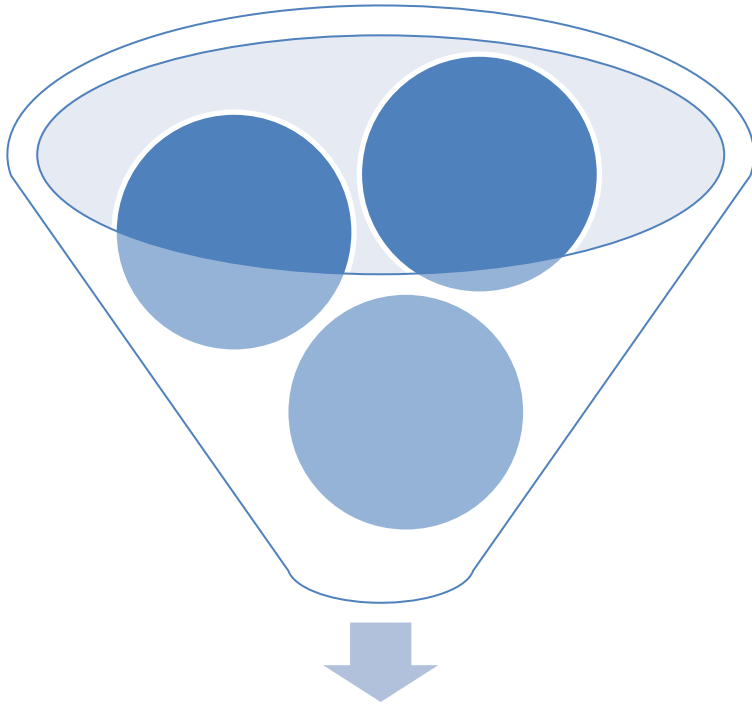


Layers communicate
using efficiently coded
Key Knowledge over
Physical Bearer

Finding and encoding
Key Knowledge is hard

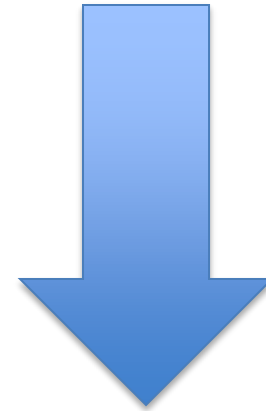


Service Management Experience



Correlated Events

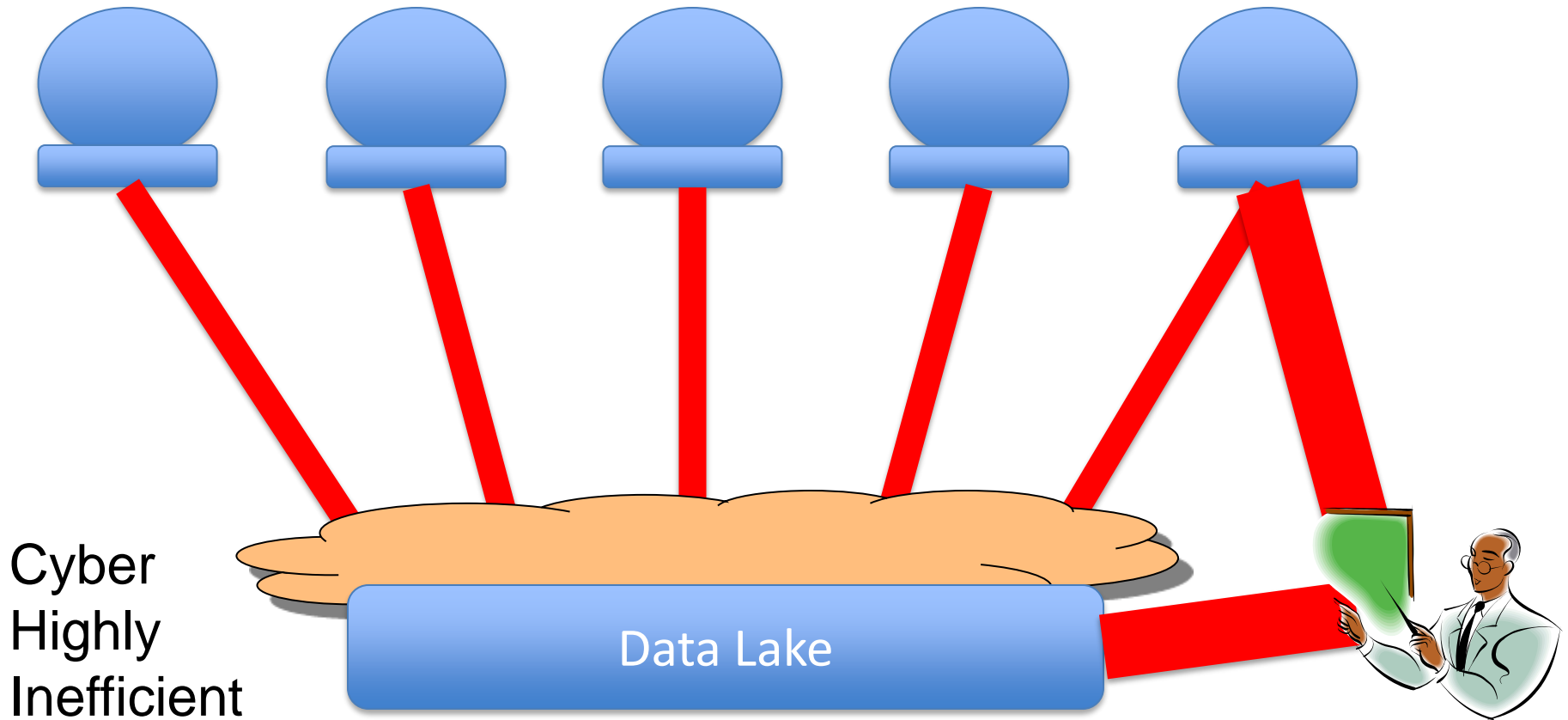
**Correlation, Rollup,
Summarisation**



**Service Affecting
Information Only**



But in Cyber why is this normal?



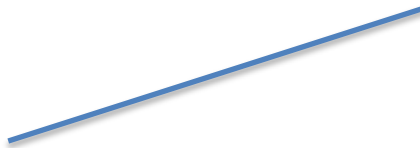


The future

Explosion of
Cyber Data



Practicality of
Centralised
Data



Availability of
Skilled
Analysts





The Challenge

- Can we use statistical analysis to improve how we find the trigger anomalies which cause an analyst to investigate?
- Can we use statistical analysis to leave more data at the edge?
- Can we use statistical techniques to delete the morass of historical data which we collected, but actually tells us nothing?
- Can we use statistical techniques to provide a model of behaviour inherent in the data, without having to keep the large volumes at all?





THANK YOU FOR YOUR ATTENTION