Presented at the 2009 ISPA/SCEA Joint Annual Conference and Training Workshop - www.iceaaonline.com

Simple Parametric Model for Estimating Development (RDT&E) Cost for Large-Scale Systems

Robert R. Jones

Paul Hardin

Anna Irvine

Research Sponsor

Mr. Robert Flowe, Office of the Deputy Under-Secretary of Defense, Acquisition Technology and Logistics, Software Engineering & System Assurance

Technomics

The Science of Informed Decision Making

Outline

1. Notional Problem Statement
2. Research History
3. Data
4. Analytical Hypothesis
5. Analysis / Results
6. Conclusions
7. Backups

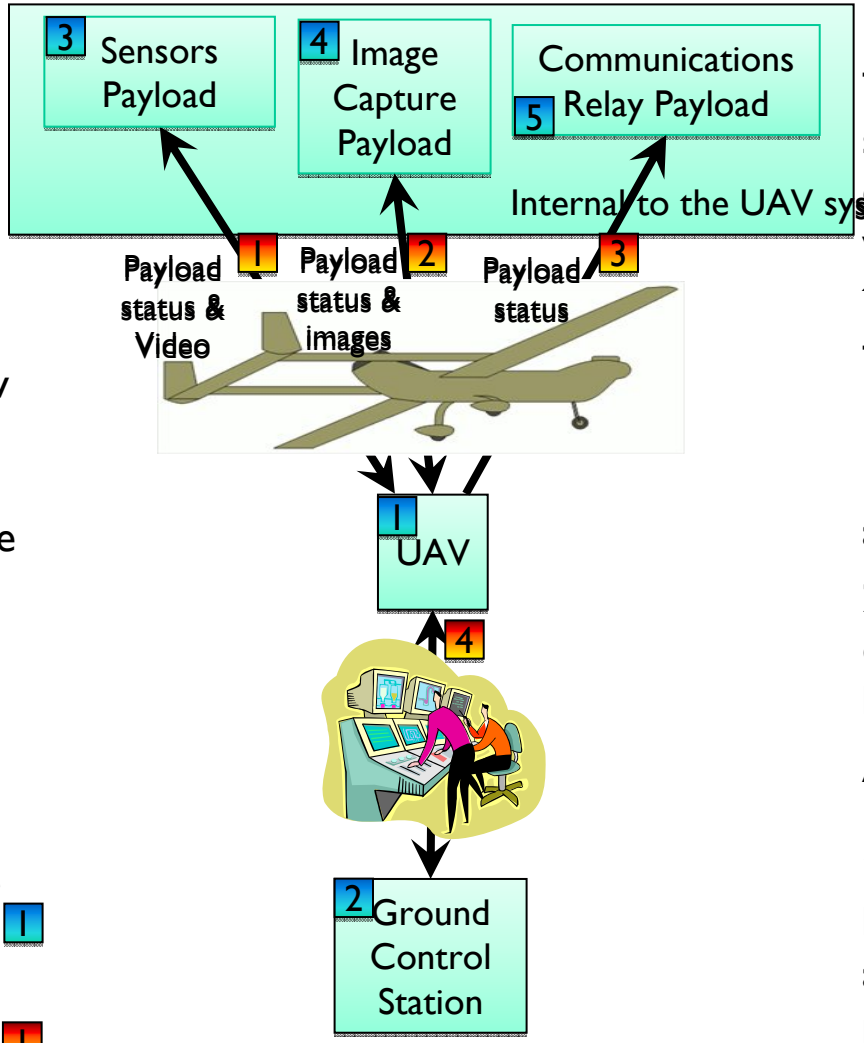
Notional Problem Statement

To set the stage for our cost research, consider a notional cost estimating problem that we believe is typical for modern-day systems.

When first describing a new system, it is not unusual to understate the degree of interdependence that will be required. For instance, the problem might initially look like this.

Unmanned Air Vehicle with

1. five new, primary components, i.e., Nodes
2. four communication transfer paths, i.e., Links



This appears to be a straight-forward cost estimating problem, but, with time, the definition of the system "expands."

The number of Nodes and Links increases because

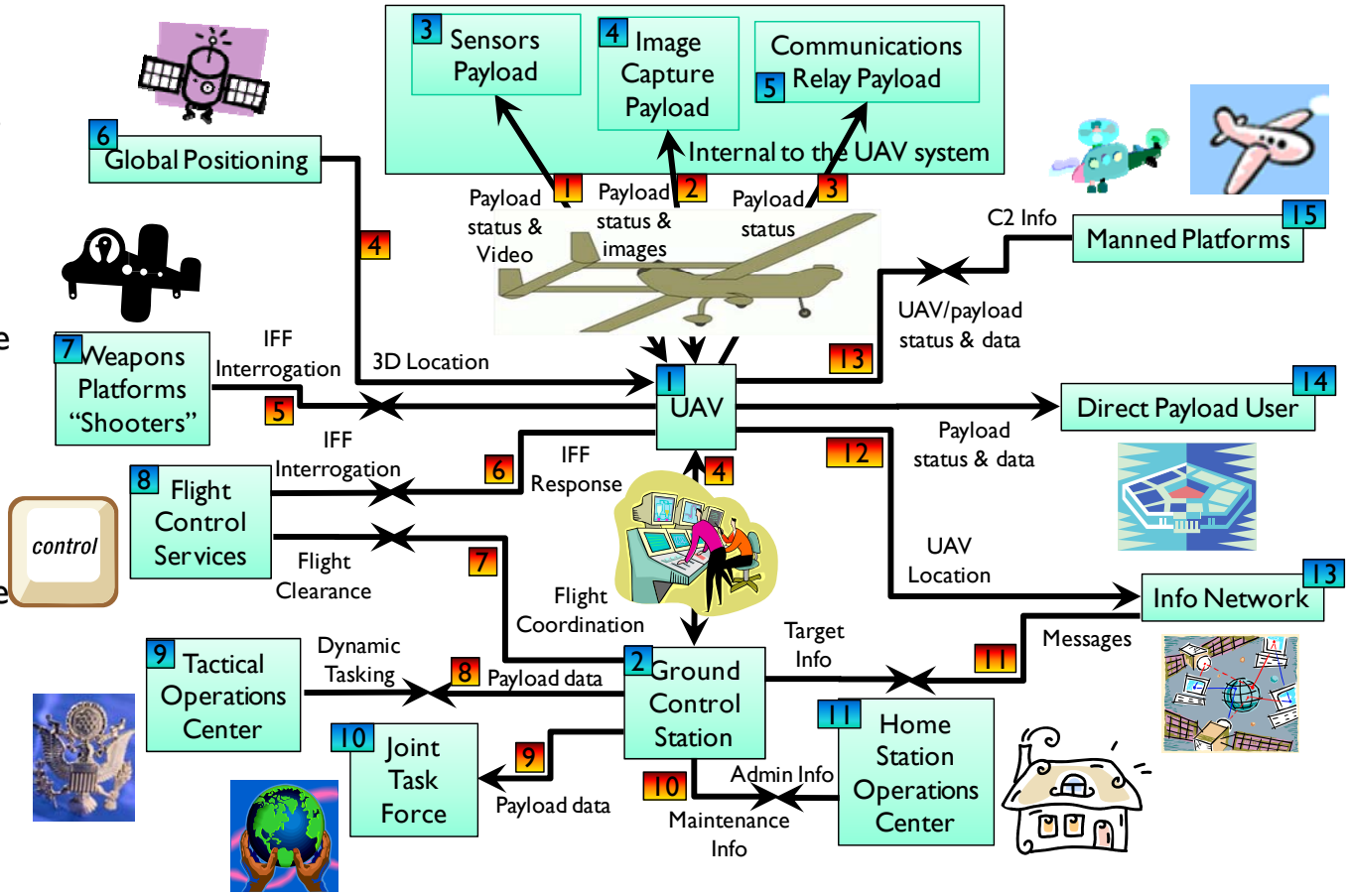
1. additional sources of data are needed by the UAV, and
2. other systems require or desire the data the UAV will be able to provide.

As a result, interoperability requirements increase, the interdependence picture becomes more complex, and the cost estimating problem becomes significantly more difficult.

Interdependence Complexity

The current, expanded UAV system is

1. at least fifteen primary Nodes (some Nodes such as "Manned Platforms" might represent several distinct Nodes, and we believe more Nodes will be identified as the system definition is refined)
2. five new Nodes, six legacy Nodes, and four "future legacy" Nodes
3. at least thirteen Links (see the Nodes comment above),
4. five uni-directional Links and eight bi-directional Links
5. two security levels, four media types, several throughput rates, etc., etc.



This is becoming a more and more difficult problem; we are confronted with increasing scope and uncertainty of the system's interdependence requirements.

Research History

Objective:

To understand the behavior of Development cost (RDT&E $) as a function of the complexity of a system and its interdependencies with other systems; improve our ability to estimate RDT&E $ and avoid cost growth.

Our research was initiated in 2003 and is still ongoing; we have adjusted our focus and emphasis to address sponsor issues and concerns.

Currently, we are sponsored by Robert Flowe, ODUSD(A&T)/SSA

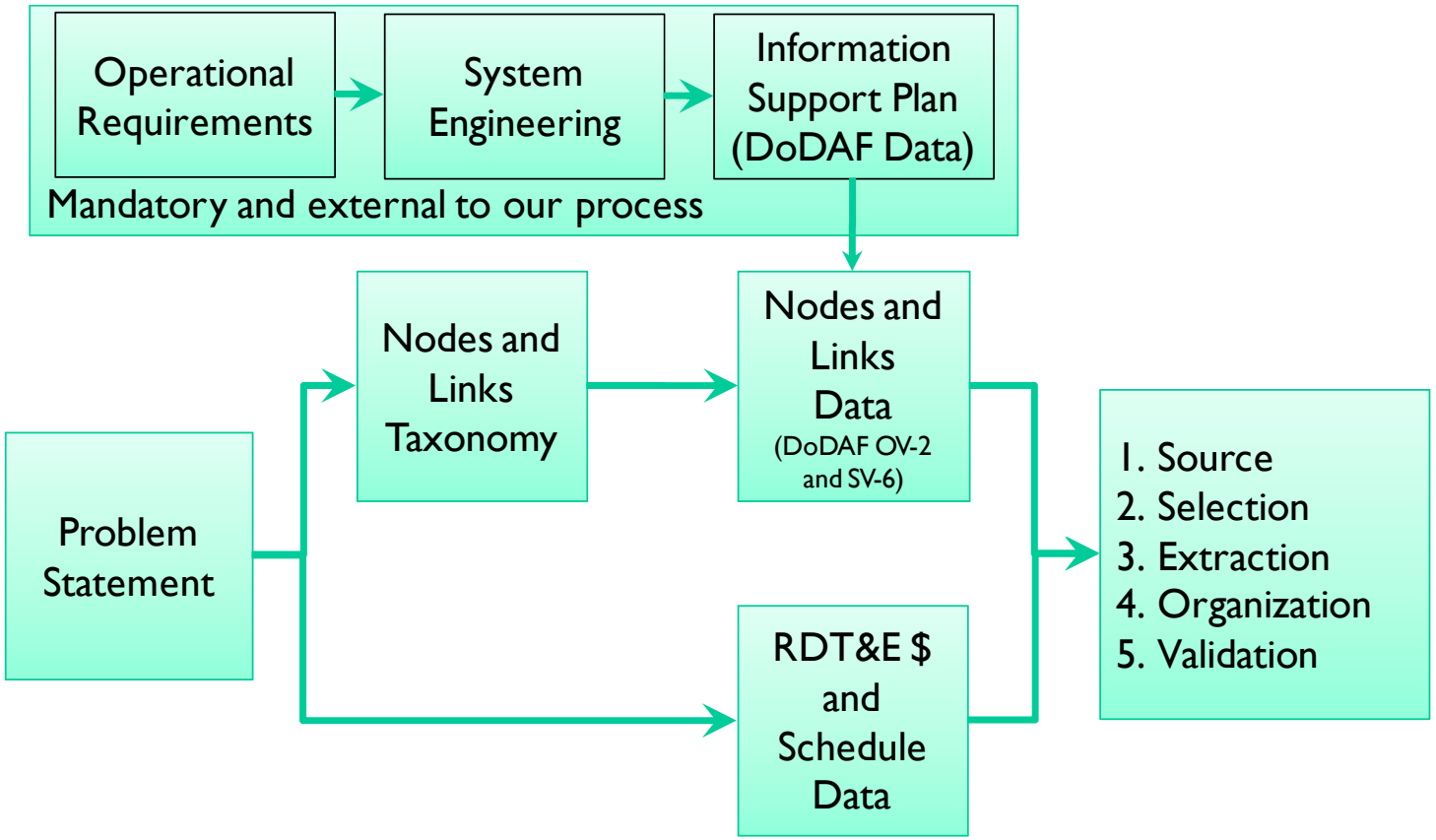
Other sponsors who have participated in the research are:

- Office of the Deputy Assistant Secretary of the Army for Cost and Economics
- Air Force Cost Analysis Agency
- Office of the Secretary of Defense, Program Analysis and Evaluation

Our principal investigators are

- Robert Flowe, ODUSD(A&T)/SSE
- Dr. Maureen Brown, University of North Carolina
- Dr. David Zubrow and James McCurley, Software Engineering Institute
- Robert Jones, Paul Hardin, and Michael Jeffers, Jr., and Anna Irvine, Technomics, Inc.
- Gary Eiserman and Ajay Choudhary, Raytheon Virtual Technology Corporation

Data

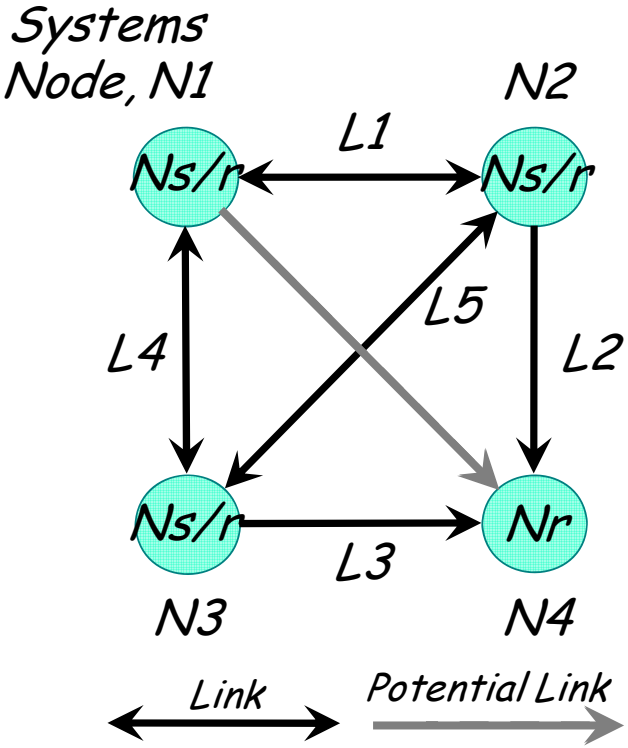


The next several slides explain the data we use, our sources, our specific data selections, our data extraction processes, how we organize our data, and how we validate our data.

But, first, we need to explain Nodes and Links.

Nodes and Links Taxonomy

Nodes and Links Taxonomy



Item	Value
Systems Nodes, *N*	
Send/Receive, *Ns/r*	3
Send, *Ns*	0
Receive, *Nr*	1
Total, *Nt*	4
Links, *L*	
Uni- directional, *Lud*	2
Bi- directional, *Lbd*	3
Total, *Lt*	5
Metcalfe Number, *LtMax*	6
Integration Density, *Lt/LtMax*	5/6
Links per Node, *Lt/Nt*	5/4

Nodes, N = an element of architecture that produces, consumes, or processes data.

Send / Receive Nodes, Ns/r = A Node that both sends and receives information.

Send Nodes, Ns = A Node that sends information.

Receive Nodes, Nr = A Node that receives information.

Links, L = A representation of the physical realization of connectivity between Nodes.

Uni-Directional Links, Lud = A Link with a uni-directional information flow.

Bi-Directional Links, Lbd = A Link with a bi-directional information flow.

Nodes and Links Data

Nodes and Links Data

1. *Source*: Information Support Plans (ISPs), Capability Development Document (CDD), and the Capability Production Document (CPD), specifically data provided by Department of Defense Architecture Framework (DoDAF) products. (See Slides 9 and 10)

2. *Selection*: DoDAF System View 6 (SV-6) and Operational View 2 (OV-2), supplemented with other views. (See Slide 11)

3. *Extraction*: Via a specialized MS Excel worksheet. (See Slide 12)

4. *Organization:* Database of thirty-three+ ACAT I, II, and III programs along with SV-6 extracted and calculated data.

5. *Validation:* Via consistency checks across the DoDAF views and "Integration Density" analysis. (See Slide 13)

(Please see the backup slides for a bit more information on DoDAF.)

Nodes and Links Data

Sources of Data

Our primary sources of DoDAF data, i.e., the ISP, the CDD, and the CPD, are *required (repeated) documents* for all programs.

Some DoDAF data advantages are:

1.use of standard, suggested templates,

2.directly-relatable to requirements,

3.products of systems engineering processes,

4.provide integrated, multi-dimensional views, and

5.enable cross-program comparisons.

And, software packages are available to insure consistency of the DoDAF data and products.

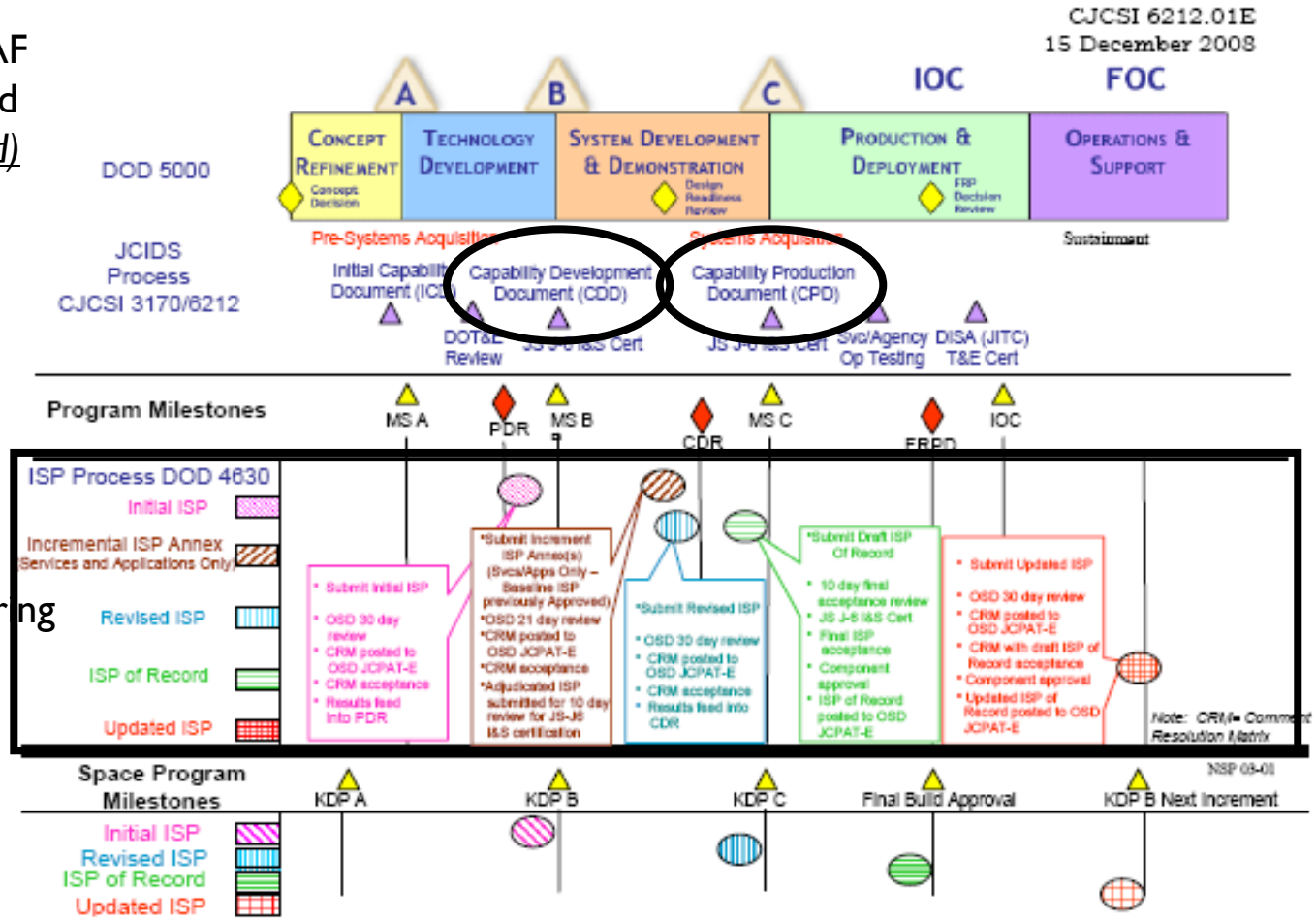


Figure B-1. Relationship between the DOD acquisition, JCIDS, and I&S Certification processes

B-5

Enclosure B

Nodes and Links Data

DoDAF Products vs Data Sources

The ISP, the CDD, and the CPD provide an assortment of DoDAF products that reflect an integrated picture of operational and system capabilities.

The data we use are provided by five of the DoDAF products. (See Slide 11)

		DoDAF Products														
Policy Source	Product	AV-1	AV-2	OV-1	OV-2	OV-3	OV-4	OV-5	OV-6C	SV-1	SV-2	SV-4	SV-5	SV-6	SV-10	TV-1
DODI 4630.8																
	ISP	X	1	X	X		X	X	X			X	X	X		X
	ISP NR-KPP	X			X		X	X	X			X	X	X		X
CJCSM 3170.01																
	ICD			X												
	CDD	X			X		X	X	X			X	X	X		2
	CPD	X			X		X	X	X			X	X	X		3
	CRD			4		4		4								
CJCSI 6212.01																
	ICD			X												
	CDD-NR	X			X		X	X	X			X	X	X		X
	CPD NR-KPP	X			X		X	X	X			X	X	X		X
	CRD (I-KPP)			4		4										
	CRD (NR-KPP)			4				4								
DoDAF																
	Integrated Architecture	X	X		X	X		X		X						X

Source: Defense Acquisition Guidebook, Table 7.2.8.1 Policy-Based Architecture Product Requirements

X = Required architecture product

1 = Acronym list

2 = Draft Information Technology (IT) Standards Profile generated by DoD IT Standards Registry (DISR)

3 = Final IT Standards Profile generated by DoD IT Standards Registry (DISR)

4 = Required for legacy Capstone requirements Documents and Capstone Requirements Document updates directed by the Joint Requirements Oversight Council

ISP - Information Support Plan (Replaces C4I Support Plan - C4ISP)

ICD - Initial Capabilities Document CDD - Capability Development Document

CPD - Capability Production Document CRD - Capstone Requirements Document

NR = Net-Ready KPP = Key Performance Parameter I = Interoperability

Policy References do not show requirements for OV-6b, OV-6a, OV-7, SV-3, SV-7, SV-8, SV-9, SV-10a, SV-10b, SV-11, or TV-2.

DODI 4630.8, Procedures for Interoperability and Supportability of Information Technology (IT) and National Security Systems (NSS)

CJCSM 3170.01, Joint Capabilities Integration and Development System

CJCSI 6212.01, Interoperability and Supportability of Information Technology and National Security Systems

DoDAF = Department of Defense Architecture Framework

Nodes and Links Data

Required / Desired DoDAF Products

DoDAF Product	General Description	Required (R) or Desired (D)	Graphical Insight	Quantity		Technical Characteristics	
				Nodes	Links	Nodes	Links
OV-1, High-Level Operational Concept Graphic	High-level graphical/tectual description of operational concept	D	X				
OV-2, Operational Node Connectivity Description	Operational nodes, connectivity, and information exchange need lines between nodes	R	X				
OV-3, Operational Information Exchange Matrix	Information exchanged between nodes and the relevant attributes of that exchange	D			X		X
SV-1, Systems Interface Description, Services Interface Description	Identification of system nodes, systems, systems items, services , and services items and their interconnections, within and between nodes	D	X				
SV-6, System Data Exchange Matrix, Service Data Exchange Matrix	Provides details of system or service or service data elements being exchanged between systems or services and the attributes of that exchange	R		X	X	X	X

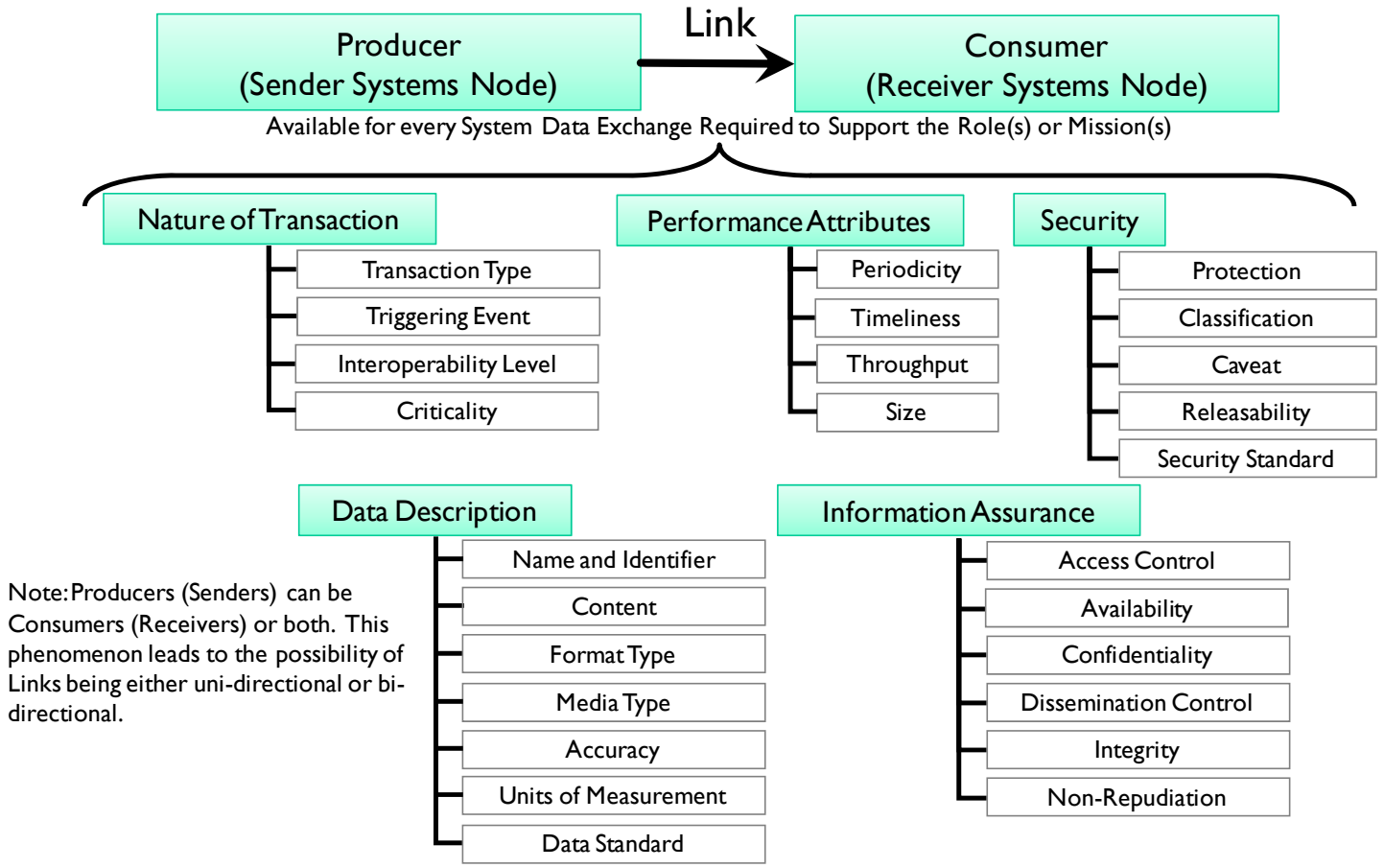
General Description Source: DoD Architecture Framework Version 1.5 Volume I: Definitions and Guidelines, 23 April 2007.

The five DoDAF products we require / desire are listed above.

The most important of these is the SV-6 (See Slide 12); it provides a tabular set of information that is readily adaptable to our specialized MS Excel worksheet used to extract the Nodes and Links data we need.

Nodes and Links Data

SV-6 Data "Template"



Here you see the DoDAF-suggested content of an SV-6.

While we monitor and extract as many of these attributes as possible, our current analysis focuses is on the numbers and types of Nodes and Links.

Nodes and Links Data

Nodes and Links Data Validation

We found an important relationship between the total number of Nodes in a system, Nt, and the "Integration Density," i.e., the ratio of the actual number of Links to the maximum possible number Links.

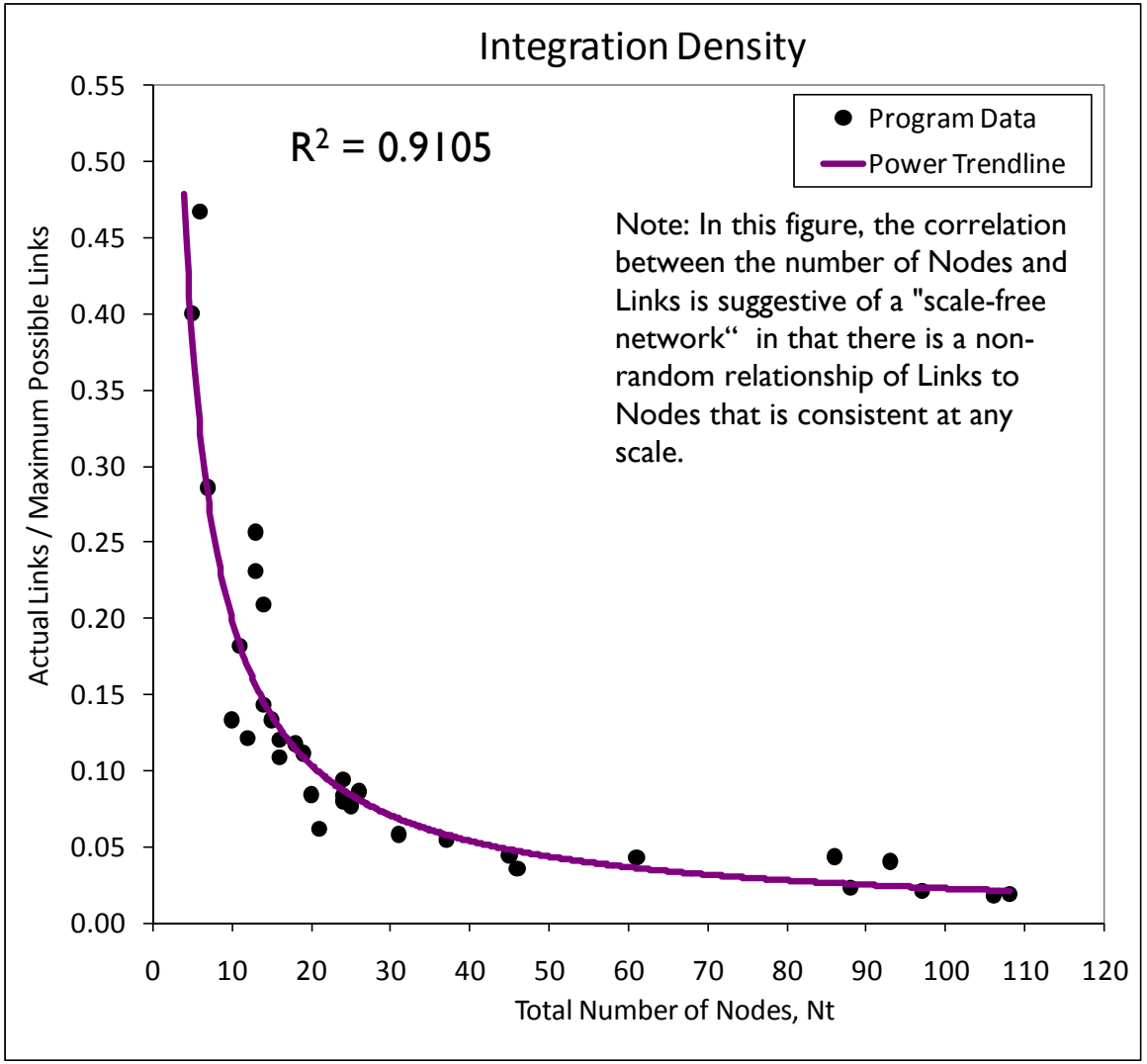
The maximum possible number of Links is given by the Metcalfe number:

$$LtMax = (Nt^2 - Nt) / 2$$

We use this relationship to "validate" extracted SV-6 data.

We also can make two important observations:

1. Given Nt, we can estimate Lt, and
2. As Nt increases, fewer and fewer potential Links are implemented.



RDT&E $ and Schedule Data

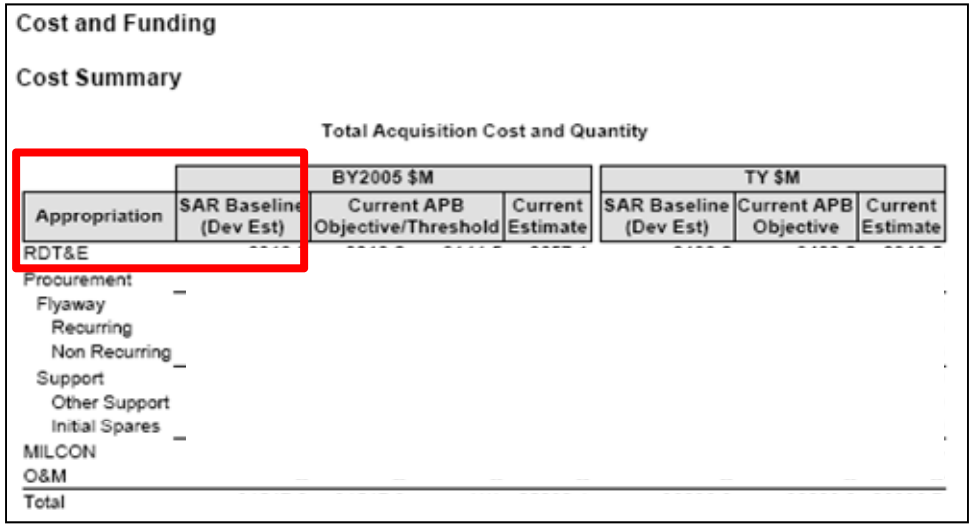
RDT&E $ and Schedule Data

1. *Source:* Selected Acquisition Reports (for ACAT I programs) and Budget Exhibits (for ACAT II and III programs)
2. *Selection (Typical December, 2007, SAR data):*

   RDT&E $: "Cost and Funding, Cost Summary, Total Acquisition Cost and Quantity, Appropriation, RDT&E, SAR Baseline (Dev Est), BY$M"

   Schedule, Milestone II or B: "Schedule, Milestones, Milestone B or II, SAR Baseline (Dev Est)"

   Schedule, Milestone C or III: Similar to Milestone B or II
3. *Extraction:* Via Defense Acquisition Management Information Retrieval (DAMIR) system or manually via Budget Exhibits
4. *Organization:* Database of ACAT I, II, and III programs along with RDT&E $ and schedule data
5. *Validation:* We use only "authoritative" $ and schedule data sources

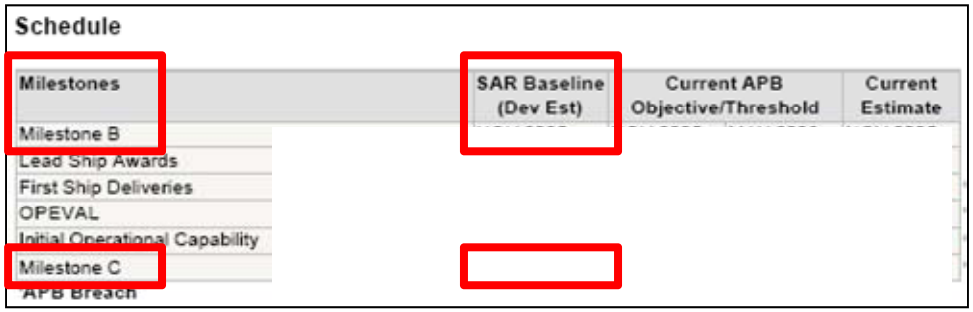


Cost and Funding

Cost Summary

Total Acquisition Cost and Quantity

Appropriation	BY2005 $M			TY $M		
	SAR Baseline (Dev Est)	Current APB Objective/Threshold	Current Estimate	SAR Baseline (Dev Est)	Current APB Objective	Current Estimate
RDT&E						
Procurement						
Flyaway						
Recurring						
Non Recurring						
Support						
Other Support						
Initial Spares						
MILCON						
O&M						
Total						

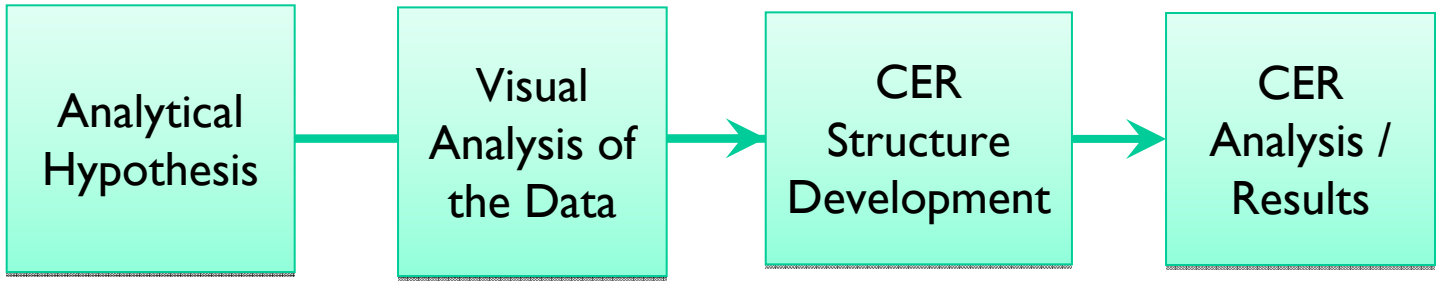


Schedule

Milestones	SAR Baseline (Dev Est)	Current APB Objective/Threshold	Current Estimate
Milestone B			
Lead Ship Awards			
First Ship Deliveries			
OPEVAL			
Initial Operational Capability			
Milestone C			
APB Breach			

As much as possible, we want the timing of our DoDAF data and our $ and schedule data to coincide.

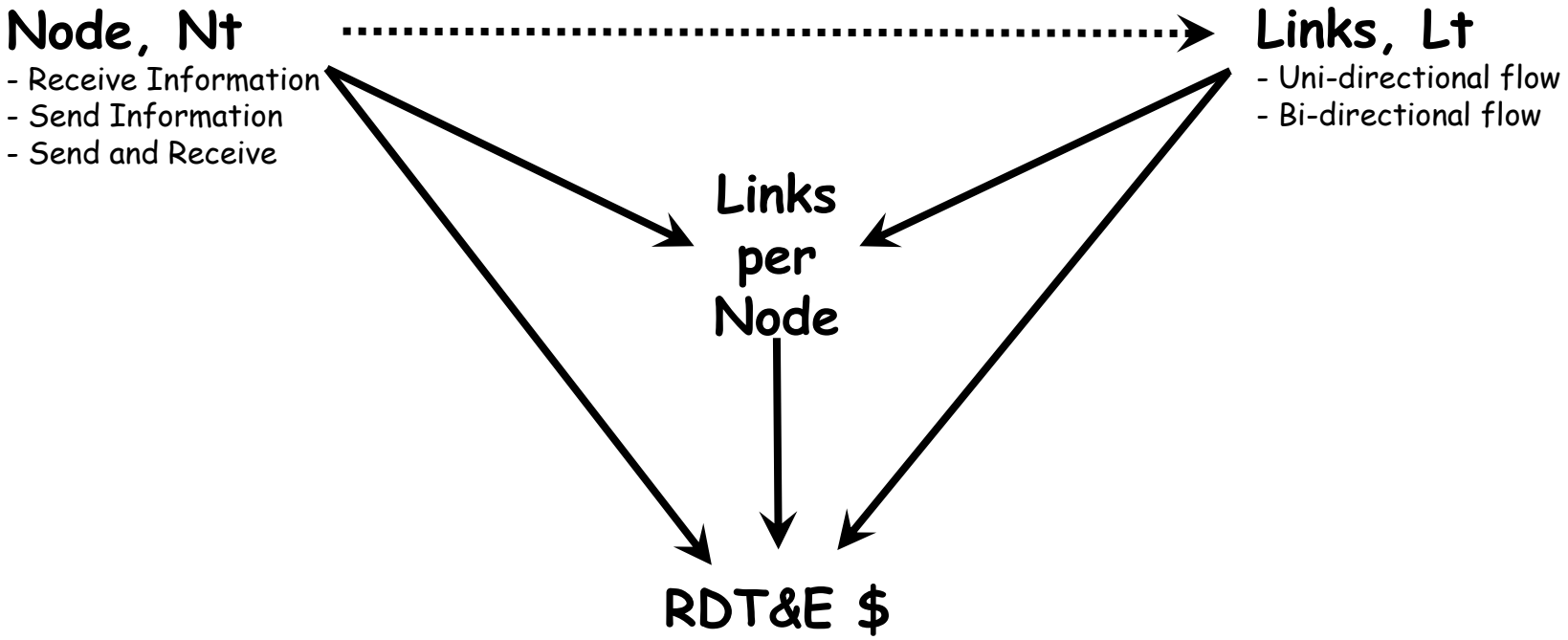
Analysis



The next several slides explain our analysis process, beginning with an influence diagram and continuing through to the final RDT&E $ Cost Estimating Relationships (CER).

Analytical Hypothesis

Analytical Hypothesis



RDT&E $ are influenced by

1. Number of Nodes and Links
2. Node and Link complexity
3. Interdependency complexity measured in number of Links per Node

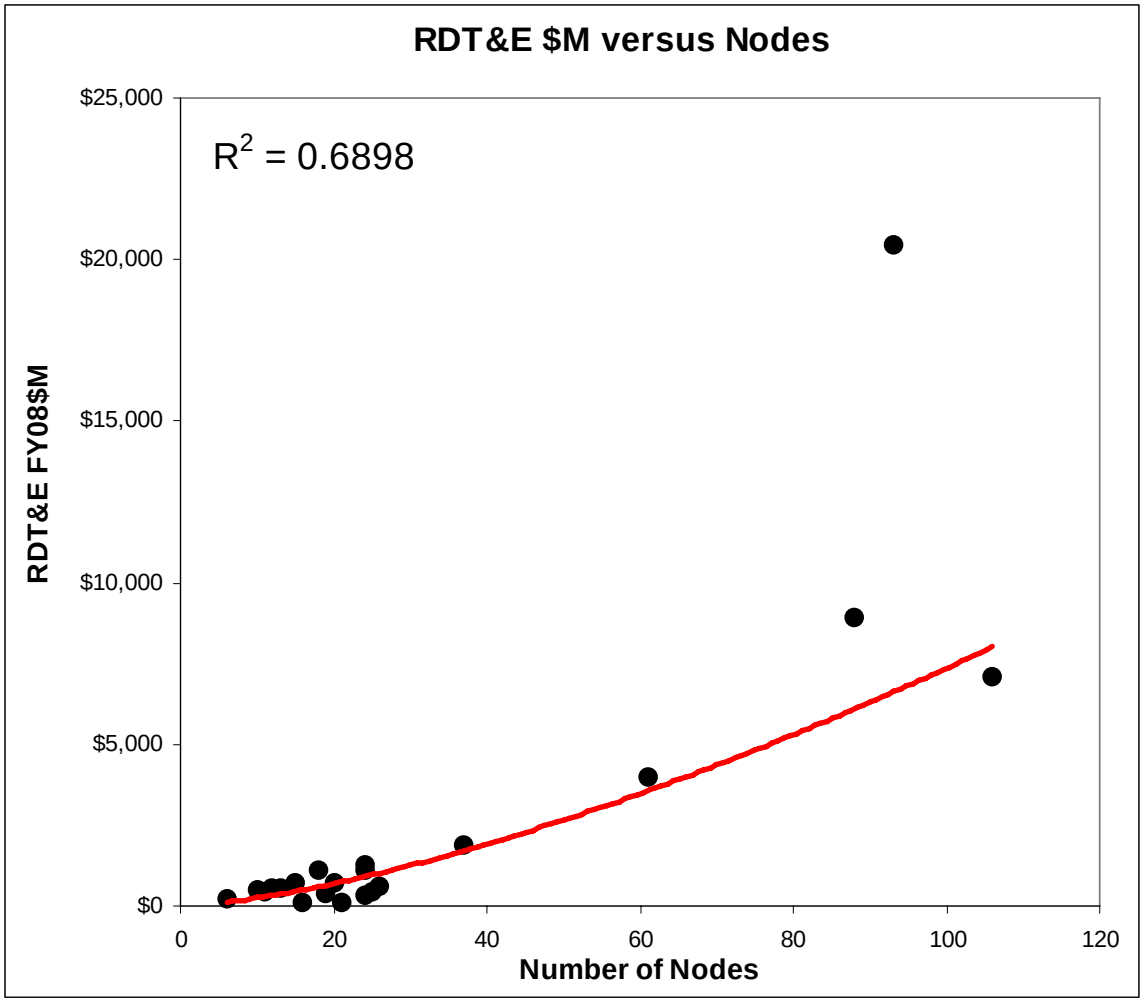
Visual Analysis of the Data

RDT&E $ versus Nodes Analysis

RDT&E $ relates to the total number of Nodes, independent of the type, or complexity.

Number of Nodes = $N_{s/r} + N_s + N_r$

In this dataset, by not considering the influence of Links in this relationship, one of the largest and most complex programs would be underestimated.



Visual Analysis of the Data

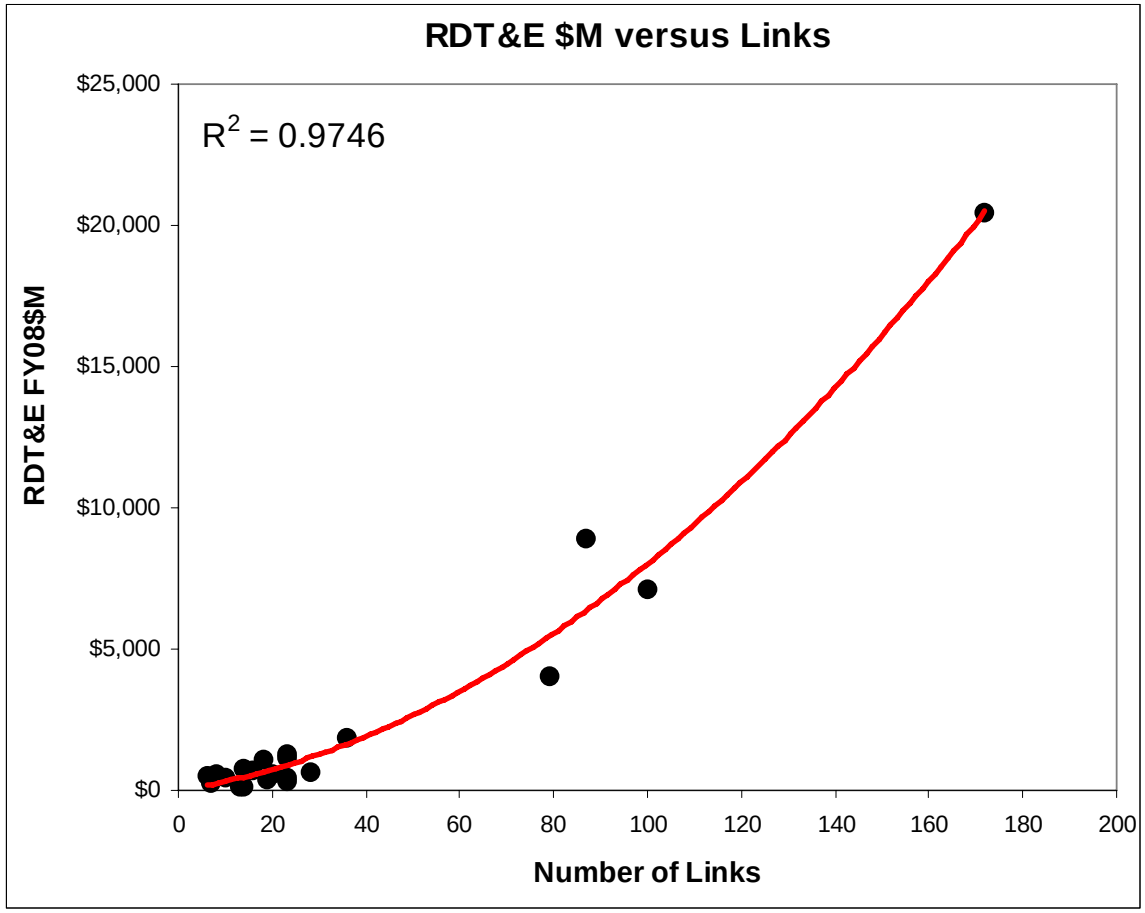
RDT&E $ versus Links Analysis

Links represent the connectivity of systems; they are a powerful, often overlooked representation of the definition of a system.

Number of Links = $L_{ud} + L_{bd}$

Although Nodes have relevance, in isolation, the number of Links appear to be even more critical than Nodes in the relationship to cost.

Understanding the relationship between both Links and Nodes is critical.



CER Structure Development

Equivalent Nodes (N_e) Analysis

$$RDT\&E\ \$ = aN_e^{\ b} \quad \text{where}$$

$$N_e = (dN_{s/r} + gN_s + hN_r)\left(\frac{L_t/N_t}{avg\left(L_t/N_t\right)}\right)^c$$

Capture complexity associated with the types of Nodes.

Capture the connectivity complexity associated with the system.

The equation parameters a, b, d, g, h, and c are estimated using MS Excel Solver and residual-minimization techniques.

CER Analysis / Results

RDT&E $ versus Equivalent Nodes

By accounting for both Nodes and Links in one equation, we improved our predictive capability over that obtained when considering Nodes and Links in isolation.

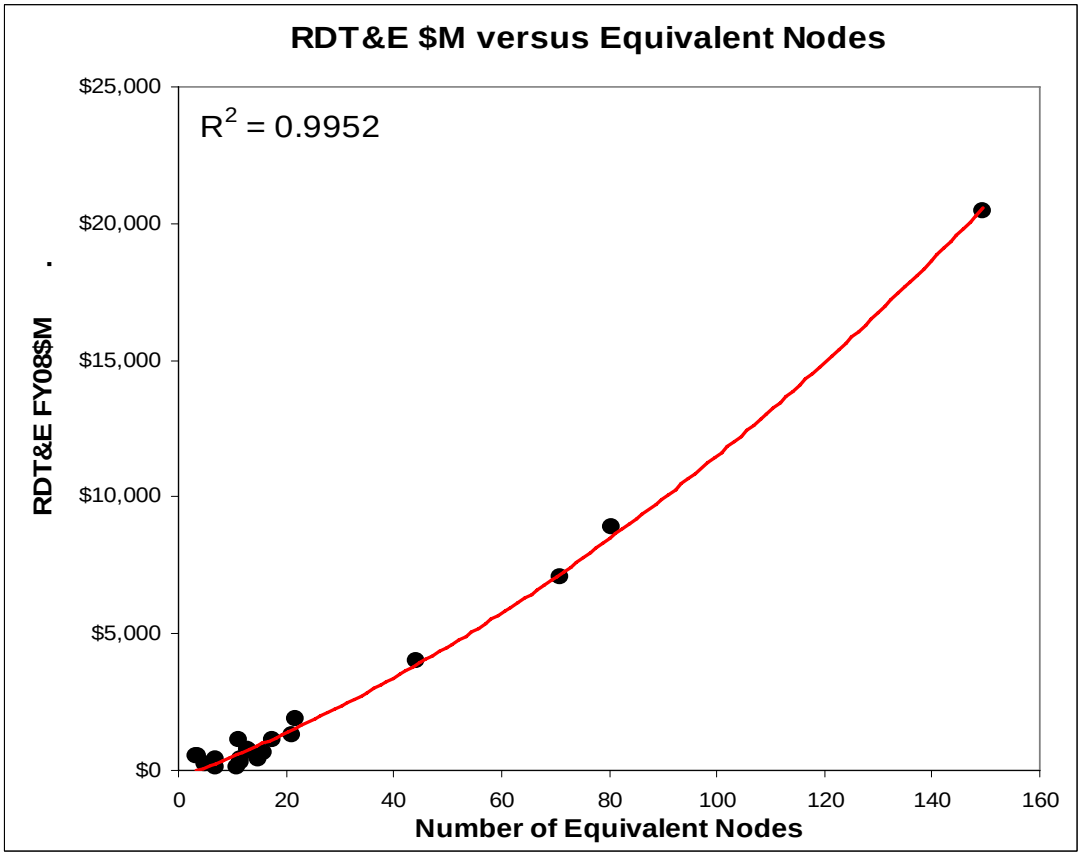
Now, our two equations are

$$(1)\ N_e = (N_{s/r} + 0.5N_s + 0.29N_r)\left(\frac{L_t / N_t}{1.02}\right)^{1.22}$$

and

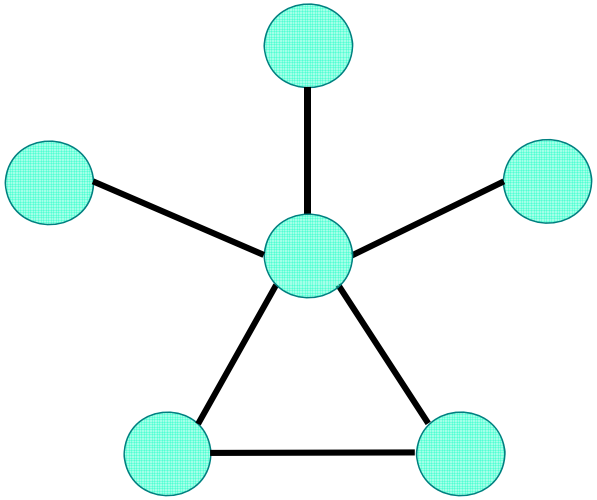
$$(2)\ RDT\&E\ FY08\$M = 20.7N_e^{1.38}$$

These two equations have explicit sensitivity to connectivity complexity and to interdependence.

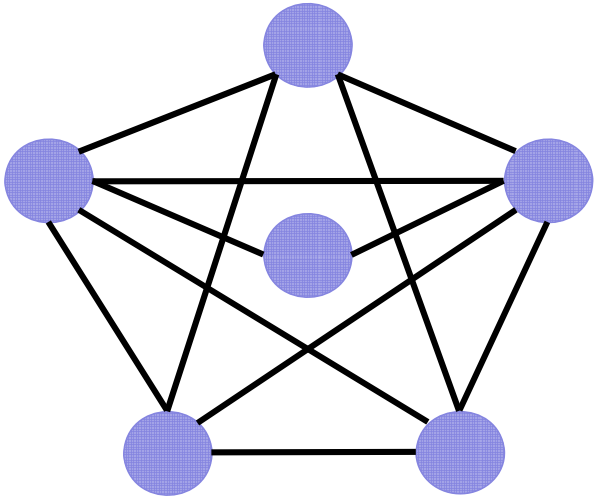


Nodes and Links Data

One Last Thought on Lt / Nt



Hub-Centric
Lt / Nt = 6/6 = 1.0
Lt / LtMax = 6/15 = 0.40



Not Hub-Centric
Lt / Nt = 12/6 = 2.0
Lt / Lt Max = 12/15 = 0.80

Our equivalent Nodes equation shows that avg(Lt / Nt) is equal to 1.02; the average system in our database can be characterized as being "hub-centric."

Why? Perhaps, (a) on-average it is true, or (b) the DoDAF products encourage systems to view themselves as "the center of the universe."

However, we have data for systems which are more complex and interdependent where Lt / Nt ~ 2; due to the power-relationships in Equations (1) and (2), these systems will have *significantly higher* development costs than hub-centric systems.

Conclusions

1. Documents normally produced during the DoD acquisition process, such as ISPs, CDDs, CPDs, provide incredibly useful information.
2. DoDAF products can be a data engine for applied-cost research.
3. Based on our analysis, Equations (1) and (2) can estimate RDT&E $ with explicit sensitivity to interdependence requirements.
4. Equations (1) and (2) can also be used for marginal RDT&E $ estimates:
   a) adding new Nodes and Links to an existing system, or
   b) combining two or more existing systems.
5. Future research will attempt to include additional parameters and address identified issues like
   a) other Node and Link characteristics,
   b) sensitivity to Legacy versus New Nodes and Links, and/or to Internal (to the Program) versus External Nodes and Links,
   c) RDT&E $ growth from MS B to MS C, and
   d) programmatic effects.

Back-Up

Definitions, 1 of 2

Node: An element of architecture that produces, consumes, or processes data. We use Systems Nodes, i.e., nodes with the identification and allocation of resources (e.g., platforms, units, facilities, and locations) required to implement specific roles and missions. We determine the number unique nodes in a system and group them into three categories, those that receive information, those that send information, and those that send and receive information. (DoD Architecture Framework Version 1.5 Volume I: Definitions and Guidelines, 23 April 2007.)

Link: A representation of the physical realization of connectivity between Systems Nodes. We determine the number unique links in a system and group them into two categories, those that have a uni-directional information flow and those that have a bi-directional information flow. (DoD Architecture Framework Version 1.5 Volume I: Definitions and Guidelines, 23 April 2007.)

"Integration Density," Lt/LtMax = the actual number of unique Links versus the maximum possible number of unique Links. We use this as a metric for the degree of integration, i.e., sharing of information, within a system relative to the maximum possible amount.

Integration: 1) The act of putting together, as the final End Item, the various Components of a system. 2) The gathering and joining together of all of the technical and functional activities and Interfaces required to link and operate the many facets of a complex weapon system. (http://www.sceaonline.net/)

We think of integration as a set of activities that are performed to achieve interoperability between and among systems. Estimates can be made of the cost associated with integration activities.

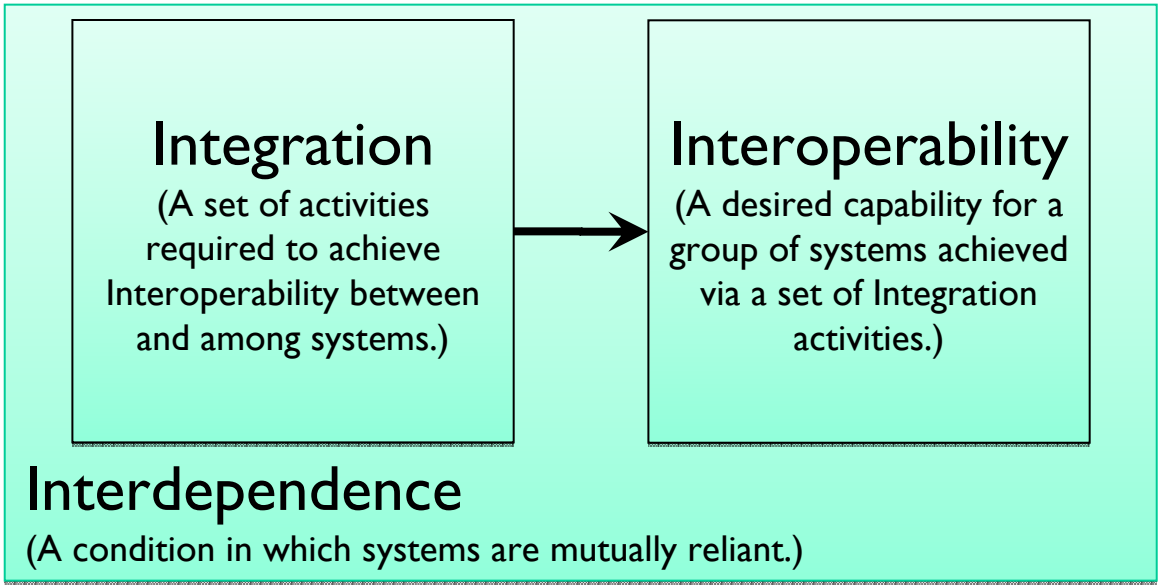
Interdependence: The output of one organization becomes one of the input for others and vice versa; organizational boundaries become less distinguishable, and the combined performance of the organizations requires complex forms of coordination. (Credit to Commentary, "Fostering Joint Logistics Interdependence," by Colonel Christopher R. Paparone, Army Logistician, Professional Bulletin of United States Army Logistics, PB 700-05-1 Volume 37, January-February 2005)

We think of interdependence as a condition in which systems are mutually reliant.

Interoperability: The ability of systems, units or forces to provide data, information, materiel and services to and accept the same from other systems, units or forces and to use the data, information, materiel and services so exchanged to enable them to operate effectively together. IT and NSS interoperability includes both the technical exchange of information and the operational effectiveness of that exchanged information as required for mission accomplishment. Interoperability is more than just information exchange. It includes systems, processes, procedures, organizations, and missions over the lifecycle and must be balanced with IA. (CJCSI 6212.01E, 15 December 2008)

We think of interoperability as a desired capability for a group of systems achieved via a set of Integration activities.

Definitions, 2 of 2



The three terms, Integration, Interoperability, and Interdependence are important to this research, so we offer the above interpretation.

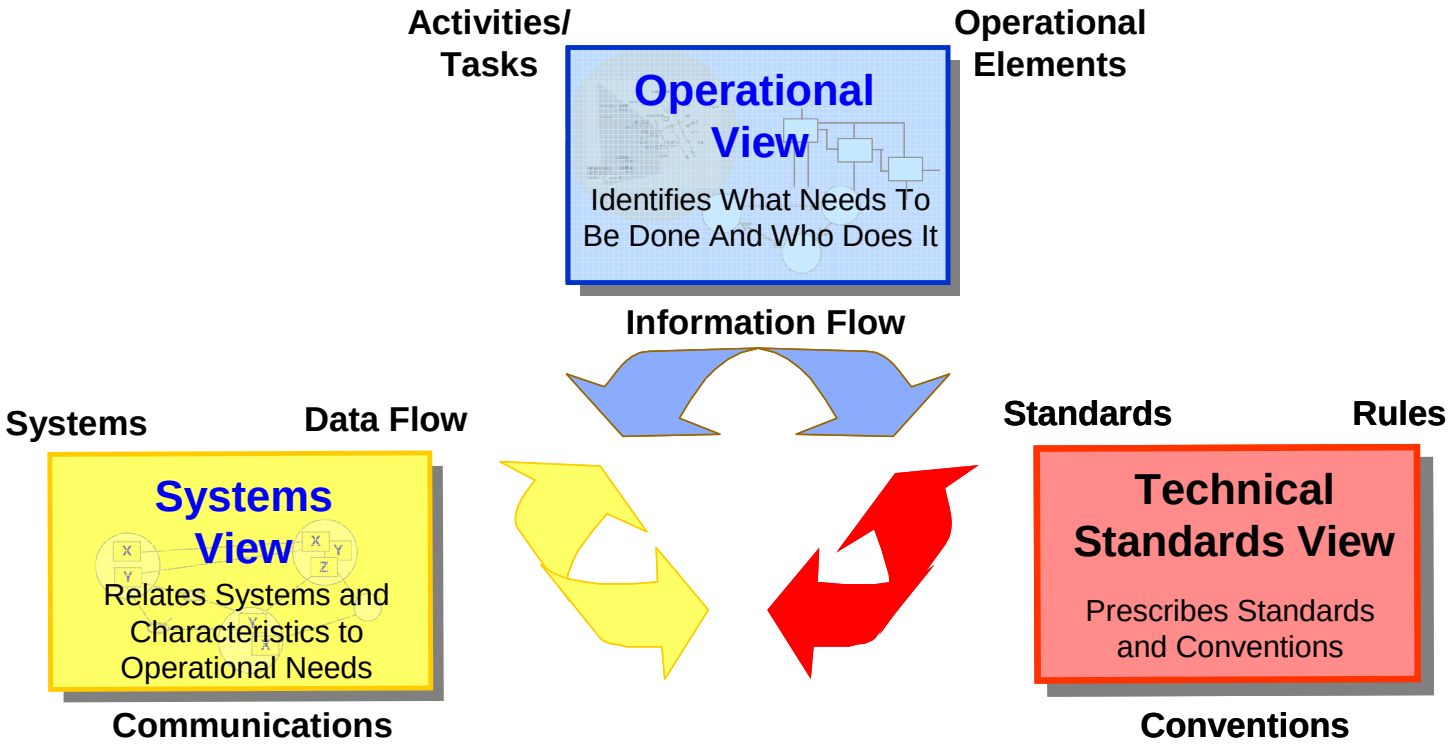
Interdependence is an end-state or condition of mutually reliant systems.

Achieving Interoperability as a capability achieved via integration activities; SV-6 data helps to define the participating systems and the nature of their connectivity.

Integration has a cost-flavor in that estimates can be made of the costs associated with the integration activities required to achieve interoperability.

Nodes and Links Data

Linkages Among DoDAF Views



DoD Architecture Framework Version 1.5 Volume I: Definitions and Guidelines, 23 April 2007.

Nodes and Links Data

DoDAF v1.5 Architecture Products, 1 of 2

Applicable View	Framework Product	Framework Product Name	Net-Centric Extension	General Description
All View	AV-1	Overview and Summary Information	✓	Scope, purpose, intended users, environment depicted, analytical findings
All View	AV-2	Integrated Dictionary	✓	Architecture data repository with definitions of all terms used in all products
Operational	OV-1	High-Level Operational Concept Graphic	✓	High-level graphical/textual description of operational concept
Operational	OV-2	Operational Node Connectivity Description	✓	Operational nodes, connectivity, and information exchange need lines between nodes
Operational	OV-3	Operational Information Exchange Matrix	✓	Information exchanged between nodes and the relevant attributes of that exchange
Operational	OV-4	Organizational Relationships Chart	✓	Organizational, role, or other relationships among organizations
Operational	OV-5	Operational Activity Model	✓	Capabilities, operational activities, relationships among activities, inputs, and outputs; overlays can show cost, performing nodes, or other pertinent information
Operational	OV-6a	Operational Rules Model	✓	One of three products used to describe operational activity—identifies business rules that constrain operation
Operational	OV-6b	Operational State Transition Description	✓	One of three products used to describe operational activity—identifies business process responses to events
Operational	OV-6c	Operational Event-Trace Description	✓	One of three products used to describe operational activity—traces actions in a scenario or sequence of events
Operational	OV-7	Logical Data Model	✓	Documentation of the system data requirements and structural business process rules of the Operational View
Systems and Services	SV-1	Systems Interface Description Services Interface Description	✓	Identification of systems nodes, systems, system items, services, and service items and their interconnections, within and between nodes
Systems and Services	SV-2	Systems Communications Description Services Communications Description	✓	Systems nodes, systems, system items, services, and service items and their related communications lay-downs
Systems and Services	SV-3	Systems-Systems Matrix Services-Systems Matrix Services-Services Matrix	✓	Relationships among systems and services in a given architecture; can be designed to show relationships of interest, e.g., system-type interfaces, planned vs. existing interfaces, etc.
Systems and Services	SV-4a	Systems Functionality Description		Functions performed by systems and the system data flows among system functions
Systems and Services	SV-4b	Services Functionality Description	✓	Functions performed by services and the service data flow among service functions
Systems and Services	SV-5a	Operational Activity to Systems Function Traceability Matrix		Mapping of system functions back to operational activities
Systems and Services	SV-5b	Operational Activity to Systems Traceability Matrix		Mapping of systems back to capabilities or operational activities
Systems and Services	SV-5c	Operational Activity to Services Traceability Matrix	✓	Mapping of services back to operational activities
Systems and Services	SV-6	Systems Data Exchange Matrix Services Data Exchange Matrix	✓	Provides details of system or service data elements being exchanged between systems or services and the attributes of that exchange

DoD Architecture Framework Version 1.5 Volume I: Definitions and Guidelines, 23 April 2007.

Nodes and Links Data

DoDAF v1.5 Architecture Products, 2 of 2

Applicable View	Framework Product	Framework Product Name	Net-Centric Extension	General Description
Systems and Services	SV-7	Systems Performance Parameters Matrix Services Performance Parameters Matrix	✓	Performance characteristics of Systems and Services View elements for the appropriate time frame(s)
Systems and Services	SV-8	Systems Evolution Description Services Evolution Description	✓	Planned incremental steps toward migrating a suite of systems or services to a more efficient suite, or toward evolving a current system to a future implementation
Systems and Services	SV-9	Systems Technology Forecast Services Technology Forecast	✓	Emerging technologies and software/hardware products that are expected to be available in a given set of time frames and that will affect future development of the architecture
Systems and Services	SV-10a	Systems Rules Model Services Rules Model	✓	One of three products used to describe system and service functionality—identifies constraints that are imposed on systems/services functionality due to some aspect of systems design or implementation
Systems and Services	SV-10b	Systems State Transition Description Services State Transition Description	✓	One of three products used to describe system and service functionality—identifies responses of a system/service to events
Systems and Services	SV-10c	Systems Event-Trace Description Services Event-Trace Description	✓	One of three products used to describe system or service functionality—identifies system/service-specific refinements of critical sequences of events described in the Operational View
Systems and Services	SV-11	Physical Schema	✓	Physical implementation of the Logical Data Model entities, e.g., message formats, file structures, physical schema
Technical Standards	TV-1	Technical Standards Profile	✓	Listing of standards that apply to Systems and Services View elements in a given architecture
Technical Standards	TV-2	Technical Standards Forecast		Description of emerging standards and potential impact on current Systems and Services View elements, within a set of time frames

DoD Architecture Framework Version 1.5 Volume I: Definitions and Guidelines, 23 April 2007.