National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Evolving Systems to Support the Global Information Grid

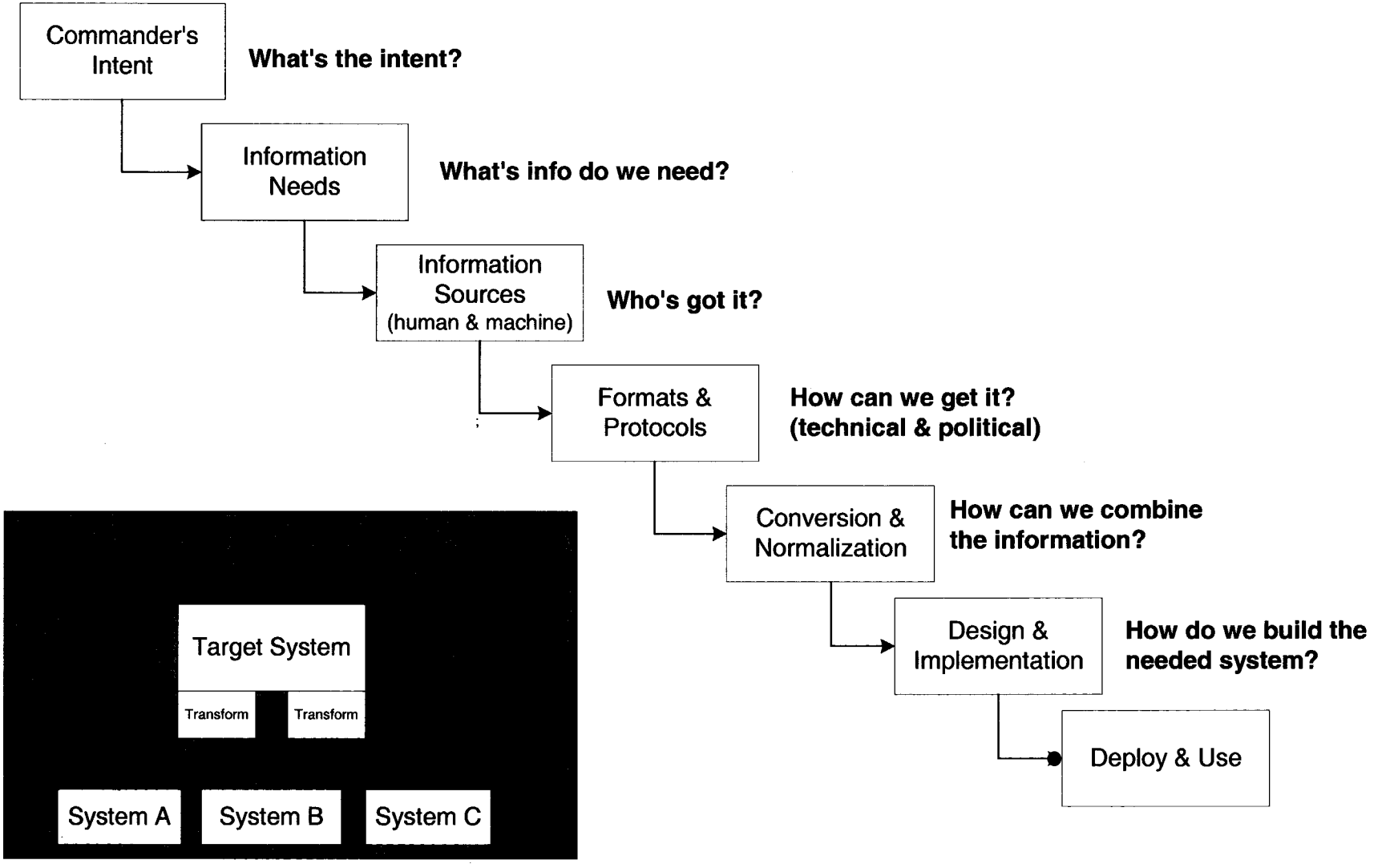
Thomas McVittie, Ph.D.
Chief Software Architect,
Deep Space Mission Systems
Interplanetary Network Directorate
Jet Propulsion Laboratory
Email: mcvittie@jpl.nasa.gov

This work supported under contracts .....

Introduction

- Information architectures have great promise for greatly enhancing our ability to understand and reason about our environments.
- Clearly, no shortage of data (information?)
- Building a true Infostructure hampered by
  - Hard to know what IS available (information hoards & fiefdoms)
  - Who has information? (people and systems)
  - Does it meet my needs? (area, subject matter, resolution, timeliness, accuracy)
  - Can I get it? (technical & Political)
  - How do I get it? (custom protocols, delivery mechanisms, organizations)
  - What does the information/data mean?
  - Can I safely combine *this* info with *that* info?
- Today – crafting of systems is primarily the domain of PEO and Integrators
  - One of the hardest parts of building any system is getting an accurate I/F specification and gaining enough perspective to understand the system's information model.
- Interesting set of capabilities coming out of DoD/OSD that will have a significant impact on our ability to build an information rich environment

Today's Idealized System Building Approach



Current Approach

- The Good:
  - Carefully orchestrated system (well understood)
  - High probability that info and presentation are suitable for user decision making process
  - Integration/combination of info from various systems is reasonably well thought through
  - We know the information that each group of users is accessing (collaboration)
- The Not-So-Good:
  - Can be VERY difficult to get sufficient information concerning other systems (formats, protocols, and meaning). Also difficult to get any sort of service commitment or guarantee of stability from systems you don't directly fund.
  - Periodic system generation – doesn't adapt readily to new information types and sources
  - Users tend to be limited to the information "we" thought they needed (not the information they think they need now)
  - Whether the system is a success or not largely depends on how well the integrator anticipates the user's needs, negotiates information availability from others (pre-arranged agreements), and predicts the physical environment.

- Lots of experience in recent efforts as to just how hard it is to really orchestrate this.

Recent DoD Initiatives

- Global Information Grid (**GIG**) and Transformational Communications (**TC**) initiatives promise to significantly transform our information (data?) collection and distribution approaches.
- TC – high capacity orbital and regional network/communications services
  - Expect exponential growth in available data to all levels.
- GIG – makes all information (data?) available to anyone who needs it
  - Targets information fiefdoms.
  - New approach is: **T**ask, **P**ost, **P**rocess, **U**se
  - Everyone publishes (posts) their information (both raw & later processed) & makes it available via an agreed upon/published mechanism.
  - Consumers can search for or subscribe to information that meets their needs.
    - Process, publish and/or use.
  - Focus is on giving power/control to the information consumer rather than the information producer.
- Similar to what has enabled the WWW, but if we're not careful, we could experience the same fate!

Infrastructure Services

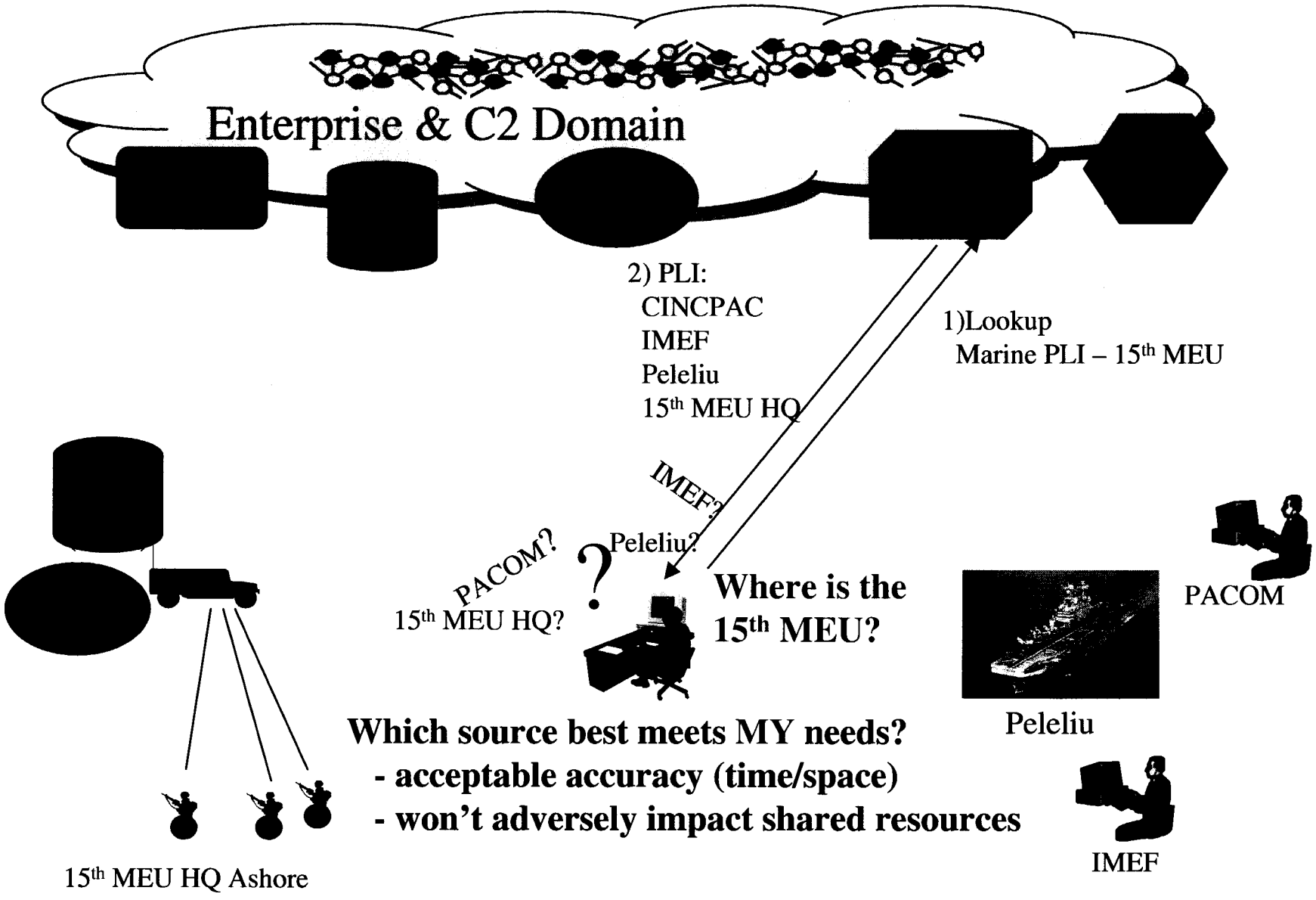
- Requires an set of widely deployed infrastructure services
- Requires a set of agreed upon approaches to defining information & meta-information

	Info providers register their capabilities, interfaces, and data schema. Consumers search for appropriate providers	UDDI, WSDL
	User & System authentication and authorization.	SAML, PKI
	Repository for information to be widely shared (proxy for actual servers)	
	Tools for finding the right information for the right need.	
	Tools for translating between domain-specific systems	

The GiG Approach

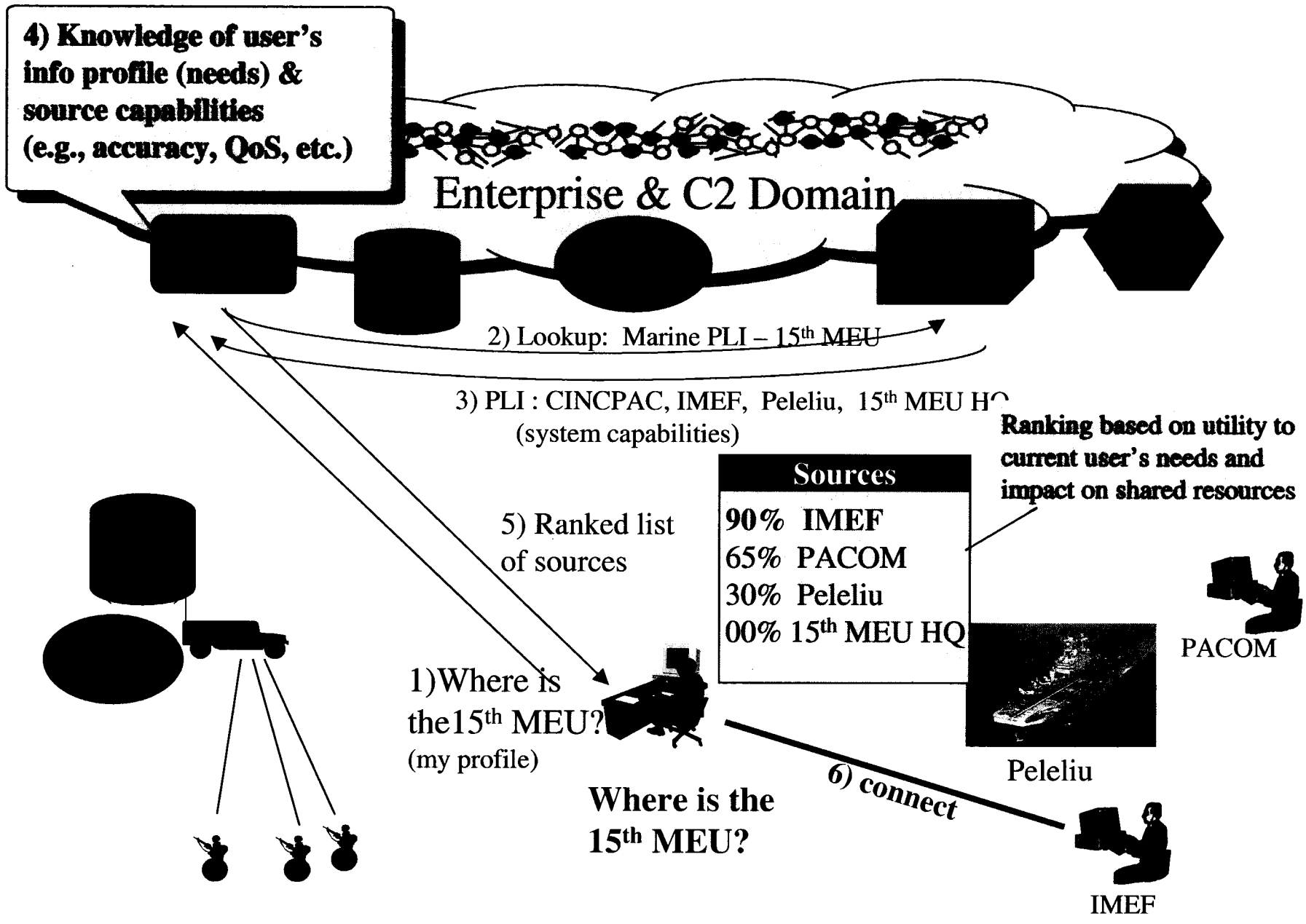
- The Good:
  - Much wider access to information (data?) for both humans & agents
  - User determines what they need.
  - Encourages Integration on the fly
    - Users can dynamically discover and connect to new information sources
      - Predator video, new intelligence reports, etc.
  - Focus is on collaboration.
- The Not-So-Good:
  - Encourages integration on the fly
    - How do we determine:
      - What information/data means?
      - Whether sources are trustworthy or authoritative?
      - Is the server able to give me the needed QoS?
      - Whether multiple data sources can be combined?
      - Whether collaborators are using complimentary sources?

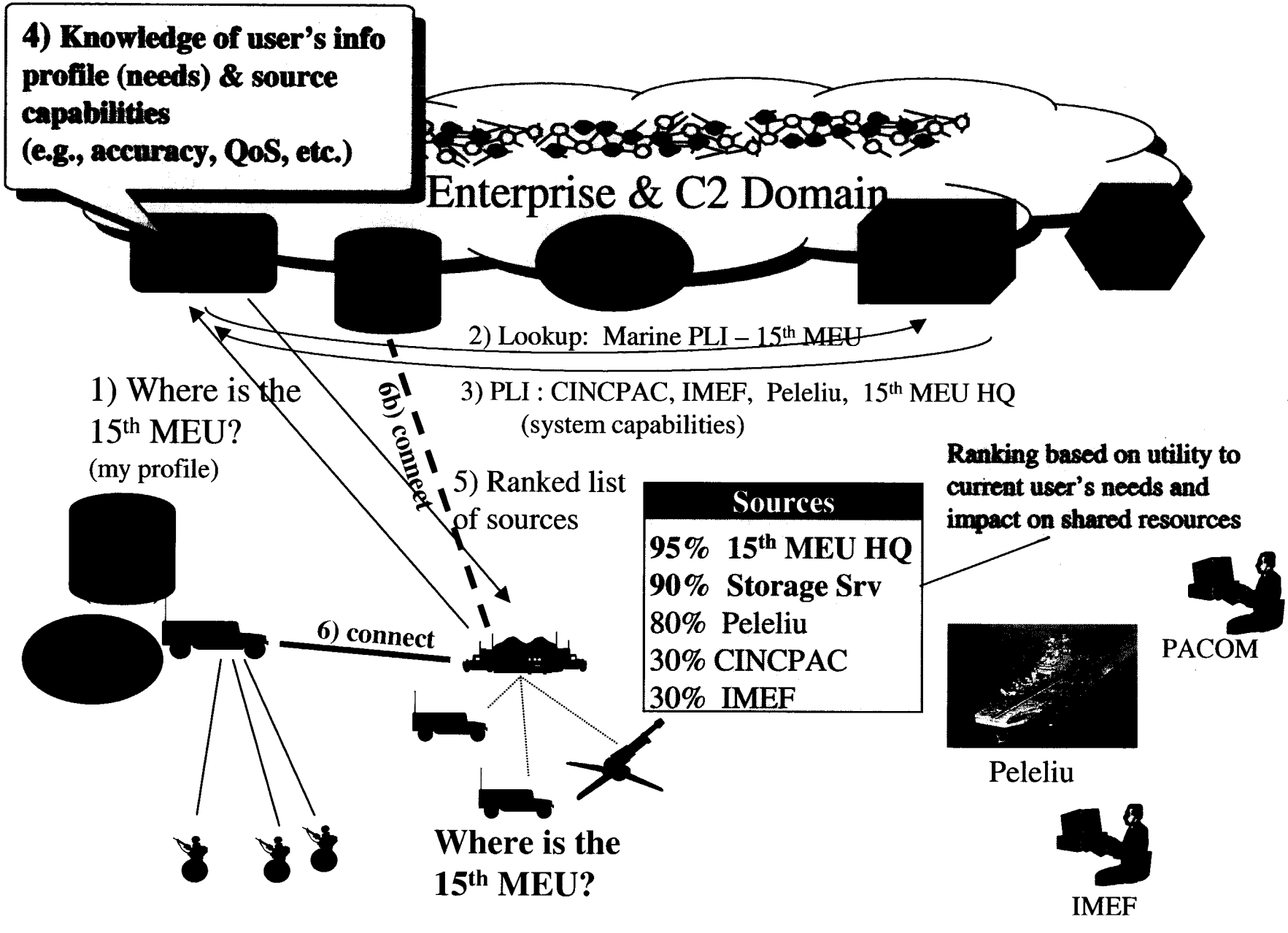
Information Discovery



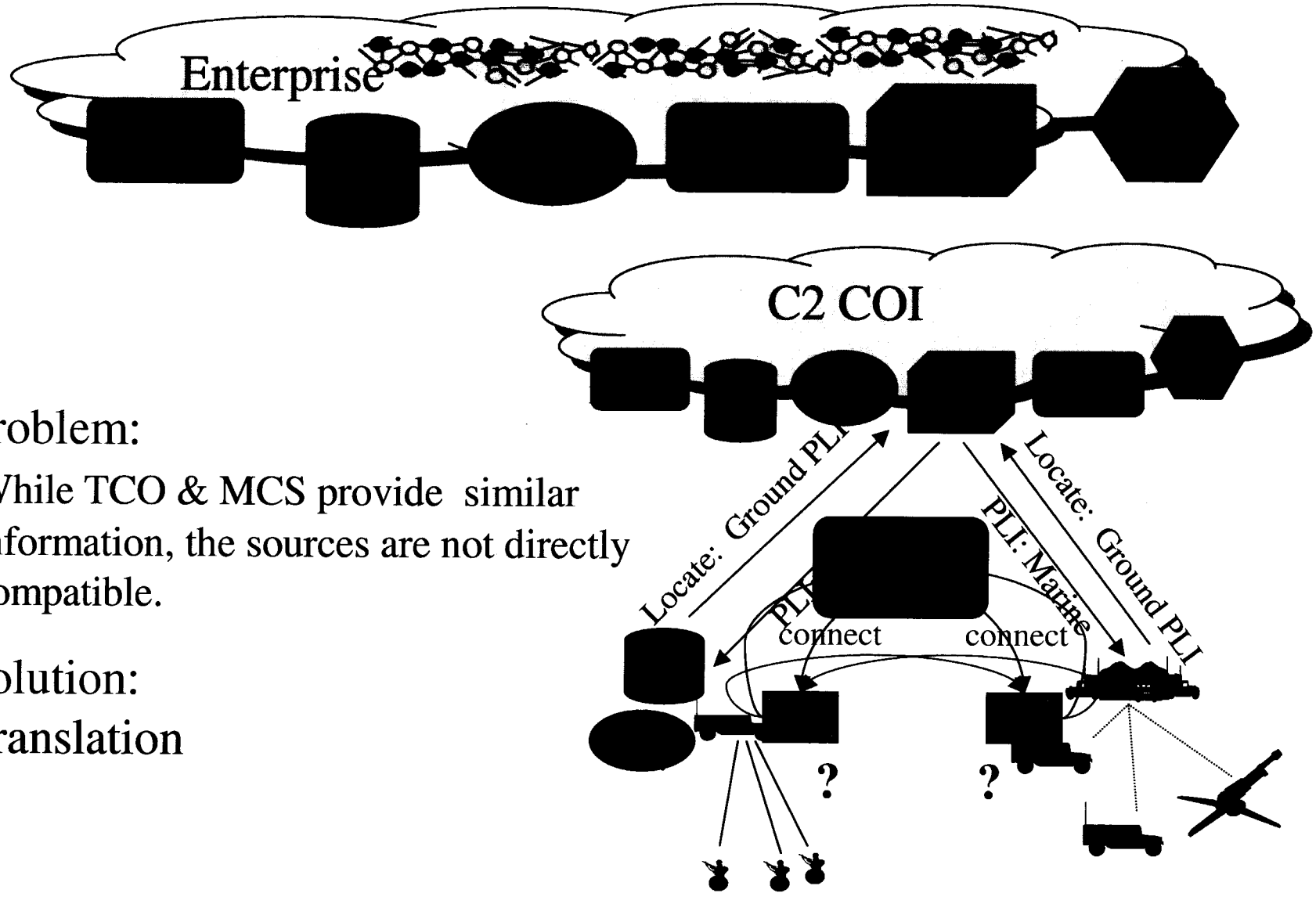
What Do We Need?

- Clear concept of information, info sources and capabilities
- Intelligent Agents (mediators) between servers & clients to provide value added processing.
  - An approach to determining which information source will best meet my needs when dealing with a potentially large number of potential candidates.
  - Mediators will need to have high levels of information concerning subject domains, users and groups (e.g., preferred sources).
- Industry Equivalents:
  - Community taxonomies (Consortia and UDDI)
  - More accurate returns on web searches.
  - Tailored information delivery
  - Users buying this book also bought ……


4) Knowledge of user's info profile (needs) & source capabilities (e.g., accuracy, QoS, etc.)
Enterprise & C2 Domain
2) Lookup: Marine PLI – 15th MEU
3) PLI : CINCPAC, IMEF, Peleliu, 15th MEU HQ
(system capabilities)
Ranking based on utility to current user's needs and impact on shared resources
Sources
90% IMEF
65% PACOM
30% Peleliu
00% 15th MEU HQ
5) Ranked list of sources
1)Where is the15th MEU?
(my profile)
Where is the 15th MEU?
6) connect
PACOM
Peleliu
IMEF


4) Knowledge of user's info profile (needs) & source capabilities (e.g., accuracy, QoS, etc.)
Enterprise & C2 Domain
2) Lookup: Marine PLI – 15th MEU
1) Where is the 15th MEU?
(my profile)
6b) connect
3) PLI : CINCPAC, IMEF, Peleliu, 15th MEU HQ
(system capabilities)
5) Ranked list of sources
Sources
95% 15th MEU HQ
90% Storage Srv
80% Peleliu
30% CINCPAC
30% IMEF
Ranking based on utility to current user's needs and impact on shared resources
6) connect
PACOM
Peleliu
IMEF
Where is the 15th MEU?

Information Exchange



Problem:

While TCO & MCS provide similar information, the sources are not directly compatible.

Solution:
Translation

What Do We Need?

- Ability to exchange data between two systems with different data schemas (aka Data Mediation)
  - Translation can be more complex than "field-to-field mappings"
    - May require prior state information (what has happened up to this point – e.g., what messages have I seen before).
    - Danger of Semantic Mismatch – Be sure you really understand the concepts not just the fields
    - Data from multiple sources (e.g., table joins)
  - Translation should be to and from a COI-standardized representation, not be between systems (NxM)
- <u>Intra</u>-COI Translation
  - translation of similar, but not identical, information between components within the same COI. Or between a COI and a closely related external partner (e.g., US exchange with Coalition Forces)
  - Less important if a COI standardizes its information representation
- <u>Inter</u>-COI Translation
  - translation of similar, but not identical, information between COIs (e.g., readiness info: coalition, US forces, maneuver, fires, logistic)

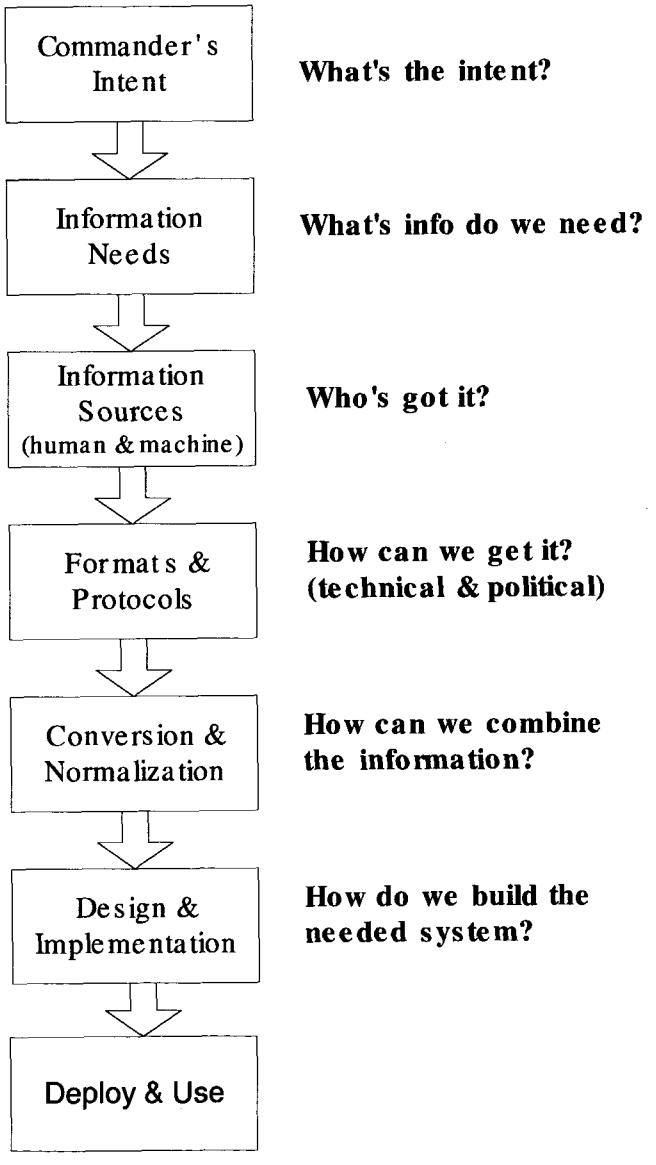
Where's the Impact

Well known location and mechanism for describing information providers & systems

Registries for interfaces and taxonomies
- Potentially eventually ontologies as well

Potential for ontology-based translation

Biggest impact is targeted towards users!



What does this all mean?

- Good longer range vision
  - Lots of promise for enhancing user's access to information & to enable meaningful collaboration.
  - Information rich source for agents
- But....
  - Very big steps
    - Changes how we operate, build & use system
- Definitely the way forward, but we need to understand the capabilities and approaches that will be needed to make this a reality.
- We welcome collaborative teaming (acadamic, industrial, etc.)
  - Thinkers/Researchers
  - Experimenters (integrate with the approaches and services)
  - System Developers
  - Deployers/Integrators
  - Users

Initial GiG Focus

- Human understandable information sources – e.g., DoD Google.
  - Better searches & access to web sites, images, documents, etc.
- Required infrastructure services
  - Security, discovery, etc.
- Machine understandable information sources - Industry Standard Mechanisms allowing current systems to define and expose their capabilities, data formats, and interfaces across networks.
    - REST, Web Services, .NET
    - UDDI and WSDL Registries
    - DoD XML Registry

[redacted]

Next Steps

- While the 1st phase of the GiG will significantly enhance our capabilities, it really sets the stage for a much more powerful set of capabilities to come.
- These are evolving and significant collaboration is welcome
- Ability to accurately characterize/represent:
  - user needs
  - info/data supplier capabilities
  - Intermediary capabilities
- Ability to match users with appropriate information sources.
- Approaches to developing and evolving Adaptable Community Ontologies
  - Ideally supporting inter-COI and intra-COI translation.
- Willingness to integrate your efforts with the GiG
  - Significant transition potential & greater leverage in enhancing access to information sources.